national STATISTICS

Studies on
Medical and
Population
Subjects
No. 67

Population projections by ethnic group

a feasibility study

Editor:
John Haskey

Assistant Editor:
Simon Huxstep

London : TSO

ISBN 0 11 621556 9

Contact points

For enquiries about this publication, contact

John Haskey

Tel: **020 7533 5121**

e-mail: **john.haskey@ons.gov.uk**

To order this publication, call The Stationery Office on **0870 600 5522**. See also back cover.

For general enquiries, contact the National Statistics Public Enquiry Service on **0845 601 3034**

(minicom: 01633 812399)

e-mail: **info@statistics.gov.uk**

fax: 01633 652747

Letters: Room DG/18,
1 Drummond Gate,
London SW1V 2QQ

You can also find National Statistics on the Internet at **www.statistics.gov.uk**

About the Office for National Statistics

The Office for National Statistics (ONS) is the government agency responsible for compiling, analysing and disseminating many of the United Kingdom's economic, social and demographic statistics, including the retail prices index, trade figures and labour market data, as well as the periodic census of the population and health statistics. The Director of ONS is also the National Statistician and the Registrar General for England and Wales, and the agency that administers the registration of births, marriages and deaths there.

A National Statistics Publication

National Statistics are produced to high professional standards set out in the National Statistics Code of Practice. They undergo regular quality assurance reviews to ensure that they meet customer needs. They are produced free from any political interference.

Contents

Foreword

In recent years, there has been a dramatic increase in the need for statistical information on the minority ethnic populations resident in Great Britain. Besides estimates of the *size* of these populations, a growing need has also been expressed for *projections* of the population by ethnic group. As a result, ONS invited expressions of interest from a wide range of experts in the field of quantitative ethnic demography to co-operate in a collaborative ONS project to investigate the feasibility of making such projections.

An expert group of 11 academic geographers, demographers and other specialist researchers was convened to consider the possibility of each separate aspect of making such projections. Their findings, collated as chapters within this volume, provide a comprehensive review of all the challenges involved, and how they might be addressed. I accept the conclusions of these findings, that projections *can* usefully be undertaken – albeit with larger uncertainty than with traditional projections – when prepared for individual ethnic communities, analysed by sex and age.

I am particularly grateful to the members of the expert group, and others who advised and participated in the exploratory workshops held to advance this feasibility review. Thanks are also due to John Haskey who co-ordinated the project, and edited this volume.

Len Cook

National Statistician

June 2002

Notes on contributors

Seeromanie Harding works at the Medical Research Council Unit in the Social and Public Health Sciences Unit at the University of Glasgow. Her main research interests lie in investigating social determinants of health and much of her work focuses on ethnic differences in health. She joined the MRC Unit just over a year ago and has been developing a programme of research in ethnicity and health, which includes international comparisons of the mortality of migrants and studies of the impact of social conditions in childhood and over the lifespan on disease prevalence in later life.

John Haskey is Head of Family Demography Unit at ONS - previously of Social Statistics Unit. He undertakes demographic analyses and estimations in the subject fields of marriage, divorce, cohabitation, adoptions and families, including one-parent families and stepfamilies. He has written on these subjects, and also on minority ethnic populations from a demographic and statistical viewpoint, in both ONS publications and other publications. He is currently President of the British Society for Population Studies and holds a Visiting Research Fellowship at Oxford University.

Michael Murphy is Professor of Demography and Convener of the Department of Social Policy at the London School of Economics and Political Science. He is Vice-president of the British Society for Population Studies. His main areas of research have been: the demography of developed societies; family and household demography; mathematical demography; population and household forecasting; kinship in contemporary societies and mechanisms for the inheritance of behaviour.

Philip Rees is Professor of Population Geography at the University of Leeds. He has produced analyses of population dynamics for three decades, developing approaches that handle inter-regional and international migration. Between 1992 and 2002 he co-ordinated the ESRC/JISC Census Programme that provides census outputs to the academic community. In April 2002, he published, with co-editors David Martin and Paul Williamson, *The Census Data System* (Wiley, Chichester), which provides essential reading for researchers wishing to use the results of the 2001 Census.

John Salt is Professor of Geography at University College London and Director of the Migration Research Unit in the Department of Geography there. His principal research interests and publications are in the field of international migration, with a particular focus on the UK and Europe. He has acted as consultant to the EU, OECD, Council of Europe, UNECE and several government departments in the UK.

Anne Scott is a Senior Research Officer at ONS. She is currently on secondment to Eurostat in Luxembourg where she is the Co-ordinator for population and social statistics for Eurostat's programme of co-operation with Mediterranean countries. At ONS she has worked on the Longitudinal Study, the development of the internal migration system and ethnicity statistics. Her PhD was on fertility and child mortality in Palestine – part of it has been published and she remains interested in the demography of this region.

Ludi Simpson is currently employed by Bradford Council and the Cathie Marsh Centre for Census and Survey Research at the University of Manchester. He has been responsible for committee and editorial work with the Local Authorities Research and Intelligence Association, the journal Statistics in Society, the Central and Local Government Information Partnership and the Radical Statistics Group. His interests lie in demography and the use of statistics by communities and governments.

Deborah Sporton is a lecturer in Population Geography at the University of Sheffield. Her research interests include the fertility of ethnic minority groups, the evolution of ethnic minority communities and, as part of the Ethnicity in the 1991 Census project, the educational participation of ethnic minorities.

Marian Storkey works as a demographer at the Greater London Authority and previously at the London Research Centre. She specialises in ethnic group demography and statistics and has written a number of publications on this subject. Her demographic experience covers works in the public, private and voluntary sectors. She has considerable experience of working on ethnic group projections having designed one of the first UK local-level ethnic group projection models for London boroughs.

Lucy Vickers is responsible for research into the production of quarterly population estimates in ONS. Prior to that, she was head of Migration Statistics Unit, ONS, responsible for the production and dissemination of internal and international migration estimates, and for analysis and research associated with those topics. She has written and presented papers on these topics.

Paul White is currently Professor of Geography at the University of Sheffield, with particular research and teaching interests in comparative European studies of ethnicity, migration and urban structures. He has also held a series of visiting posts at the Universities of Zaragoza (Spain), Cagliari (Italy) and Paris I (France). He has served in the past as Secretary and later Chair of the Population Geography Research Group of the Royal Geographical Society.

Introduction

John Haskey

This volume marks the culmination of the first stage of a project initiated by the Office for National Statistics (ONS) to consider making projections of the population by ethnic group. A review of the needs and priorities for ONS statistical work and analyses involving ethnic group had been carried out in the late 1990s, following the successful publication of the four volumes in the *Ethnicity in the 1991 Census* series. These volumes had shown that the introduction of the ethnic group question in the 1991 Census had yielded an immense quantity of valuable data, from which important new analyses had been possible. When planning these four volumes, it had been decided that they should contain cross-sectional – rather than longitudinal – analyses, with possibly a subsequent volume being devoted to longitudinal analyses using data from the ONS Longitudinal Study (LS).

The review involved consulting Government Departments and a wide range of academic specialists – and the review was advertised, inviting expressions of interest. A questionnaire was used, which listed a number of possible avenues for future ONS analytical work on ethnicity, besides longitudinal analyses; respondents were not only invited to express their views on the desirable order of priority, but also to indicate whether they would be interested in participating and contributing to a project on any of the topics, should it be chosen.

Work on projections of the population by ethnic group was accorded the top priority, followed by an interest in longitudinal aspects – both quintessential demographic topics, but ones that also undoubtedly benefit from a multi-disciplinary approach. Given that the inclusion of an ethnic question in the 2001 Census was virtually assured – partly due to the four volumes having established its value – and because longitudinal analyses would shortly be much strengthened by the linking of 2001 Census data into

the LS, it was decided to embark on making population projections by ethnic group.

From those who had expressed interest in this project, an expert group was convened, consisting of academic demographers and geographers, and other researchers specialising in the field of what might be termed 'quantitative demographic ethnicity'. The first meeting of the expert group was held at ONS in October 1999, under the encouragement and direction of Dr Peter Goldblatt, Divisional Director of the (then) Demography and Health Division.

For this projections project, four stages were identified as being desirable, the first of which being a 'feasibility stage' to investigate whether it would be *possible* to make projections of the population by ethnic group. This feasibility stage was seen as ensuring that the appropriate groundwork, in terms of both concept and practicality, was explored and documented for issues such as the availability of appropriate data, methodology, consistency constraints and criteria, and other relevant considerations.

To launch this feasibility stage and act as a scene-setter, an overview background paper was first written, considering the needs such projections would serve – and the role they would play in informing current and future policy. In addition, the paper identified some of the more important practical and theoretical hurdles which would have to be faced and overcome. The paper concluded that, for a wide range of purposes, projections of the population by ethnic group *are* a necessary element in the armoury of available national demographic data, rather than a discretionary 'optional extra'.

For example, just as estimates of the population by ethnic group are used to monitor racial discrimination and measure equality of opportunity, so too can the corresponding projections be used to estimate the likely future proportions of the different minority ethnic populations in the overall population – to determine the monitoring targets likely to be required in the future. Such scenarios should also help the understanding of how present and future targets relate – and the fact that there need not necessarily be a steady transition from one to the other. Also, the fact that several local authorities have already each made their own set of projections is another, very tangible measure of the existing need.

Following on from this scene-setting paper, the different aspects and issues which need addressing and investigating were identified – and each member of the expert group selected one particular issue for research, producing a 'feasibility paper' on that issue. The aim was that all the challenges in making ethnic projections would be discussed and recommendations made to try to meet them. More importantly, the central question for discussion was whether the issue concerned posed a surmountable or an insurmountable obstacle to making the desired set of population projections by ethnic group.

First drafts of all these 'feasibility papers' were presented to a Workshop held at University College London on 13 June 2000; a number of other experts and researchers in the field of ethnicity were also invited to attend and contribute to the discussion (see the Appendix to this Section for participants). The purpose of the workshop was twofold: to provide ideas for the authors to improve their drafts, and also to ensure that the common and related themes running through the papers could be reassessed and re-evaluated by the gathered authors so that the papers could be redrafted to form a coherent set. This volume contains the 'scene-setting' paper and the set of finalised feasibility papers as chapters.

In writing their feasibility papers, members of the expert group were also requested to address the considerations involved in producing *subnational* projections of the population by ethnic group. A precise description of the areas that might be adopted was not given, since practical considerations of data availability and estimation would most likely dictate what would be feasible.

Appendix: Workshop on projections of the population by ethnic group – held at University College London on 13 June 2000

Presenters of draft feasibility papers (in order of presentation)	Other participants/discussants	
Michael Murphy	Peter Goldblatt	Marian Fitzgerald
Philip Rees	Peter Aspinall	Charlie Owen
Marian Storkey	Roger Ballard	Tony Champion
Seeromanie Harding	David Owen	Lucy Vickers
Ludi Simpson	Eileen Howes	John Haskey
Deborah Sporton	John Hollis	Simon Huxstep
Paul White	Dave King	Angela Donkin
Anne Scott	Chris Shaw	
John Salt		

This is a blank page

Overview

Overview of the issues raised in considering the feasibility of producing population projections by ethnic group

John Haskey

General

The most immediate impression gained on reading the chapters in this volume is the sheer number of considerations to be taken into account on adding an ethnic component to the 'traditional' form of projections – those by age, sex, and area. Furthermore, some of these considerations have ramifications for the other considerations, so bringing an interactive effect, thus making for increased complexity. Adding ethnicity certainly introduces three main issues which have to be catered for: first, because individuals decide their own ethnic group, it is essentially one of self-identity – which is subject to change over time; second, the ethnic group of a child is definitely not one of self-identity, but will be assigned by someone else and will be either that of the mother, or the father (if different), or some other ethnic group, including of mixed origin; third, the traditional data sources which provide information on past trends in the three main components of fertility, migration and mortality lack an ethnic dimension, so that other data sources and estimation methods have to be brought into play.

Review of ethnic projections literature

The first chapter, by Marian Storkey, provides a comprehensive review of the literature on projections by ethnic group and closely allied projections, such as those by race. The purpose of the review not only serves as an introduction and provides a broad context to the present project, but also gives an appreciation of exactly what has been achieved previously and elsewhere. Furthermore, it was hoped that the details of these different kinds

of projections would indicate the most frequently occurring problems, and the solutions adopted – which would provide valuable lessons in the current context.

Overall, the main finding is that there have *not* been an enormous number of ethnic projections made; some 20 countries, and four local authorities within England have produced such projections of varying sophistication. Although there is considerable diversity in the types of projections made – with regard to ethnic group, race, country of birth, and the characteristics of the projections, such as time span – almost invariably the cohort component method has been used. In general, few details have been given of the assumptions made, or how they were formulated, or of the methodology employed, suggesting that an ethnic dimension necessitates a more pragmatic and subjective approach – even within the framework of the tested cohort component method – than is the case with traditional projections.

Some particular findings are worth noting. The projections made by certain local government authorities may have been slightly easier to make than national ones since special local sources of data have usually been available – sources that are not available nationally. Also, local circumstances may have assisted the making of projections; for example, in Bradford the majority of the minority ethnic population is South Asian, and a names analysis distinguishes the South Asian ethnicities relatively well. However, despite their advantage with special sources of data, the main weakness of the projections made by local authorities lies in the necessary manipulation of those data to estimate the component rates by ethnic group – and this weakness will inevitably also apply, and to a greater extent, to any new national or subnational projections that are made.

A sensitive issue can be the projected *relative* sizes of different ethnic groups, especially if those sizes depend critically on the migration assumption made. The problem can be compounded by the background data on migration – and particularly its estimation by ethnic group – being the weakest among those for the different assumptions that have to be made.

The review chapter usefully includes work carried out on 'back projections'; the production of earlier estimates of minority ethnic populations from more recent ones. While these back projections may not be considered projections in the usual sense, their results are not only useful in their own right, but also generate new estimates from which to derive trends – which in turn are invaluable in formulating assumptions for the future.

The USA and Canada, and a number of European countries, have produced projections by race, nationality, ethnicity or birthplace. However, very few have made projections for *every* ethnic group, the US being one of the few exceptions. Sweden and Denmark have distinguished the foreign-born population and the native-born descendants of immigrants – which allows, on combination, the total minority ethnic population to be considered. Some countries have made projections of only a single race or ethnic group – which can simplify the methodology somewhat.

Perhaps the overall conclusion from the experience of other countries is that the cohort component method has proved most resilient by its widespread use – though this popularity might be largely due to the fact that it focuses attention on the aspects which matter most – the three sets of assumptions – even if there are large practical difficulties in formulating them. Another lesson is that the estimation and methodological details should be described much more fully – a criticism which this volume should go a considerable distance to rebut!

Consistency considerations and the possible role of microsimulation

The second chapter, by Mike Murphy, considers the characteristics of the conventional cohort component model, the multistate population projection method, and the class of microsimulation models – and the advantages for using each of these models with projections by ethnic group. Within this framework, the way in which the practical consistency requirements can be met are outlined, as well as mentioning the data input implications.

There is a strong prima facie case for using the cohort component model to produce projections of the population by ethnic group, by virtue of its particular properties. As well as producing such projections directly, the model can also be used to produce the *overall* population projections and then, separately, the *proportions* of the population belonging to the individual ethnic groups can be projected, finally multiplying the two elements together. However, whichever route is chosen, the

cohort component method is extremely demanding on input data.

One difficulty with the cohort component model springs from the fact that it operates on the population of women, adding men into the projected population via the sex ratio of births, and removing them according to male mortality rates. As a result, the model cannot incorporate factors such as partnerships between men and women – and some treatment of partnerships is necessary if account is to be taken of the fact that a child's ethnicity depends on both that of the mother and that of the father.

When the population can move between different states, the cohort component model is known as a multistate population projection method. This method is particularly appropriate to ethnic projections if one wishes to incorporate the facility to allow for change in ethnic self-identity. Such change can be treated as a special kind of migration. The multistate model when applied to regional projections is often referred to as a multi-regional model – which can also easily deal with changes in ethnic self-identity.

The dimensions of the 'state space' – the number of ethnic groups to be projected, as well as the other characteristics such as age, sex and possibly area in addition – imply that the data demands are very large. For example, with six ethnic groups, and using single years of age up to age 100, the total number of transitions per year easily reaches 6,000.

Microsimulation models hold their own attractions, being much more flexible than the cohort component models and the multistate methods. For example, given that transitions occur to individuals, each person's ethnicity is derived by applying pre-specified probabilities, and so the relationships between the ethnicities of children and those of their parents can be easily represented. Nevertheless, these models are very demanding of computing resources and background information from which to estimate the required probabilities. In addition, being stochastic, the models produce different results each time.

While a microsimulation model may provide an intellectually more appealing probabilistic approach to change and also carry the advantage of having an in-built measure of the *variability* of the projection results – it is perhaps more appropriate as a research tool than as providing one particular reference set of projections. Overall, it is concluded that a cohort component model in which the population can move between different states – i.e. a multistate model – could, in principle, provide the basis for ethnic projections – which appears to be the more practical proposition.

New models for projecting ethnic populations

The third chapter, by Phil Rees, continues the theme of the previous one by developing the cohort component model for the particular application to projections by ethnic group in the United Kingdom. While the treatment is essentially theoretical/algebraic, the model is formulated with the practical situation very much in mind. The existing official national projection model is first expanded to one catering for every ethnic group, by simply repeating the model for each ethnic group. Then, immigration and emigration are allowed for separately, rather than as a single net effect. Then, inter-country, or inter-regional, migration is separately distinguished – which results in the model becoming a multi-regional projection model. In the next few steps, the model is further developed to cater for: sets of subnational areas (recognising that there are distinct areas where the majority of the minority ethnic populations live); changes in ethnic self-identification; and children's ethnicity in relation to that of their parents.

The algebraic formulation for the successive steps of the developed model is succinct (although the reader will have to cope with, or skip over, according to preference, the many subscripts and superscripts), and, with each additional term introduced, is usefully complemented by a discussion of the practical considerations concerning that feature. For example, when ethnicity is added, it is pointed out that the number and nature of the ethnic groups to be used partly depend on how well the related fertility, mortality and migration rates can be estimated; the broad ethnic groups used in the 2001 Census brings the advantage of simplicity, albeit with some associated loss of homogeneity. The author concludes that a more detailed classification, based on the 2001 Census ethnic question, 16 groups in all, is more appropriate and should be used.

The different sources of data from which to estimate fertility and mortality rates for this

classification are then identified and discussed, recognising some of the difficulties, such as return migration biasing mortality rates. With regard to migration, disaggregating it into its components – of immigration, emigration and inter-country migration – and catering for these elements separately in the model requires a reassessment of the projection method. Using migration data which records a change in location between two points in time necessitates employing *transition* probabilities in a transition model – which in turn requires the national projection model to be reformulated to incorporate, among other factors, inter-country migration.

The inter-country model is then extended to allow a larger number of subnational areas; two possible alternative sets are suggested: metropolitan areas and the regional remainders; and a set of almost 70 Local Government Areas with the largest proportions of minority ethnic populations. It is concluded that reasonable estimates of the required inputs to the projection model can be made, using data from the 2001 Census. However, the author considers it sensible to use the 5-fold ethnic group classification, rather than the 16-fold one. Finally, changes in ethnic self-identification and mixed ethnicity births are catered for, through the introduction of conditional probabilities.

Perhaps the most important message from this chapter is that multi-regional projection models are fully established as a standard, and bring the important advantage that internal migration consistency is automatically achieved. However, the number of variables required for input into the model can be vast, although they can be estimated from more aggregated data. The key lesson is that careful preparation and justification of the inter-regional intensities of migration is necessary. The author also suggests that it would be feasible to add the dichotomy: 'born in the UK'/'born elsewhere' to the ethnic classification – which would enable first and second and subsequent generations of migrants to be distinguished.

Overall, the chapter provides designs for ethnic projection models at three different areal levels: national, subnational and local – and the conclusion is that robust and reliable projections are indeed feasible.

Estimating methodologies and assumptions for ethnic projections

The following chapter, by Ludi Simpson, reviews the methods and supplementary data sources that have been used by certain local authorities to estimate the change in populations by ethnic group in their areas between 1991 and 2001. The purpose is to draw lessons not only for data and methodological aspects of making national projections by ethnic group, but also for a possible common framework so that national and subnational ethnic projections can be placed on a compatible basis.

Bradford, Oldham, and Rochdale Councils, the London Research Centre, and the Centre for Research into Ethnic Relations have each estimated local authority district populations by ethnic group in 1991 and projected them forwards. Because of the differences in the ways in which these organisations have estimated the change in ethnic composition in the districts concerned, the author has helpfully summarised the data and methods used by each organisation in a couple of Tables – which not only provide a useful and compact documentation, but also allows for easier comparisons.

An important conclusion which the author draws is that it *is* possible to project the population by ethnic group for every local area – using census and other data that are available consistently across the country for such areas – and the extra data that are available in certain areas only can be used to validate and improve the reliability of the projection model results for those areas. A corollary, concluded by the author, is that the type and nature of the special data will inevitably determine which methods of analysis are most appropriate.

Some examples of the special data are given: Bradford has copied birth records from its Health Authority maternity and child health administrative systems to yield birth counts by ethnic group and sex, and by age and parity of the mother; Bradford has also kept the counts of intended first address and status of immigrants, which it has derived from notifications from Port Health Authorities to the Health Authorities of the immigrants intended first addresses; and the London Research Centre has documented 15 sources of information that improve the estimation of where refugees are living who wish to seek asylum in Britain.

The chapter next considers the task of formulating assumptions, and discusses whether the different vital rates might be expected to converge between ethnic groups. There is certainly evidence that distinct differentials *do* exist, and, in cases where appropriate data are available, that these vital rates are not always converging. The conclusion is that careful measurement is a high priority for ethnic projections in order to base assumptions on accurate information from the past.

Overall, the author recommends, as a standard method for both national and local population projections by ethnic group, a cohort component model with separate components for births, deaths and migration. He provides some detailed suggestions for the number and type of categories to be used in the projections, and recommends that subnational projections should be produced for selected districts – with ethnic proportions above 10 per cent – plus regional remainders. Also, he recommends the use of an age-specific propensity to transfer between each pair of ethnic groups, as a means of modelling mixed parenthood as well as changes in self-identity. Finally, an important contribution of the chapter is its emphasis on the necessity for the prior consultation and obtaining widespread support, especially *political* support, for the publication of projections by ethnic group.

The starting population for ethnic projections

The following chapter, also by Ludi Simpson, addresses the issue of the starting, or base, population for ethnic projections. The availability of a base population can be very easily taken for granted – or deemed of lesser interest than the type of projection model, or the challenges for estimating the rates constituting the assumptions. The chapter illustrates the importance of the starting population – and the necessity for thoroughly considering all the details connected with it.

One of the most obvious requirements of a base population from which projections by ethnic group are to be made is that it is sound – and that the ethnic classification of the base population is consistent with those of the component assumptions. Given that results from both the 1991 and 2001 Censuses will together provide information on trends, as well as recent estimates, an understanding of the comparability of the two census ethnic classifications – and how they can be used together – is also crucial.

The issues to be decided in conjunction with the starting population are listed in detail; while many are to be resolved in other chapters, they are most usefully rehearsed in this chapter. To mention three which arise on considering data from both censuses: geographical boundaries change over time, but a common set is required for comparability; the special populations of the Armed Forces and of students may need separate treatment in the projection; and adjustments are required to derive mid-year population estimates from the census population.

A straightforward comparison of the 1991 and 2001 Census results by ethnic group is not only beset by different ethnic classifications; they are also affected by the different treatment of undercounts. In the 1991 Census results, the undercount was only partially estimated, whereas in the 2001 Census results a full allowance for the undercount will be incorporated. The chapter briefly describes the considerable work that has been undertaken in adjusting the 1991 Census populations, both to make them consistent with the mid-1991 population estimates by district, and also to estimate the population by ethnic group in each ward, using the overall rate of non-response in each ward. Further work has refined this approach, by including an additional differential undercount between ethnic groups in each ward.

As a result, a better set of population estimates is now available than the published 1991 Census results, even though the sizes of the remaining errors are unknown. The author points out the important point that the uncertainty about the undercount could explain the variation in vital rates derived from the corrected census populations – and that variation should not be built into a projection.

The remaining part of the chapter considers two possible ways of comparing 1991 and 2001 census ethnic group data using two ethnic group classifications. One of them – an eight-fold one – allows a direct allocation between 1991 and 2001 categories. However, much depends on the consistency of individuals' responses, and the author suggests a probabilistic approach.

The conclusion that the reader will draw from this chapter is that improved corrected population estimates have been derived from the

1991 Census, but that there is a practical limit – perhaps now reached – to the removal of the undercount bias and the remaining uncertainty. As a result, vigilance and care will be needed to ensure comparability of both censuses' data regarding ethnic group – not only for base populations, but also for the vital rates derived from them.

Fertility

The sixth chapter, by Deborah Sporton and Paul White, investigates the subject of ethnic fertility, recognising it to be arguably *the* critical assumption in ethnic projections, given that most of the minority ethnic populations have a much younger age profile than that of the White population. We are reminded that traditional national population projections normally incorporate higher and lower fertility variants, and while it is not suggested that variant fertility assumptions should be formulated for each ethnic group, it is pointed out that variants are used to reflect the uncertainty associated with that assumption. The sensitivity of the projection results to the fertility assumption can be large – especially when too high or too low an assumed rate has a cumulative effect over the number of years of the projection period. The fertility of some minority ethnic populations can be influenced by both that of the country of origin, and that of the country of current residence.

A full discussion of the difficulties in projecting fertility by ethnicity is given, covering: the issue of ethnic self-identity; reliance on births by mother's country of birth; the ethnicity of newborn children; and inter-ethnic marriage and families. In addition, the subject of the geography of ethnic fertility is raised, suggesting that geographical variations might be expected since where minority ethnic groups have settled in concentrated numbers, cultural norms, attitudes and behaviour are more likely to be preserved. Some evidence exists for such an effect, and the authors suggest that ethnic fertility assumptions should take account of local variation.

Examining the available data sources for estimating ethnic fertility, the authors review and discuss the alternatives: the 1991 Census; the Longitudinal Study; the General Household Survey; the Labour Force Survey; Hospital Episode Statistics and Maternities data from Health Trusts; and finally birth registrations. The authors then compare the 1991 total period fertility rates (TPFRs) estimated by ethnic group (using birth registration data and the 1991 Census) with an alternative set derived from the Labour Force Survey; they conclude that the marked differences in the TPFR for some ethnic groups highlights some of the data issues previously discussed, but also conclude that there are real differentials that need to be incorporated into the fertility assumptions.

Of course, current levels of fertility are needed, but there is an even greater need for time-series data on ethnic fertility to gauge the likelihood of future changes. TPFRs by mother's country of birth have been available since 1981, and may be converted, relying on a number of assumptions, to TPFRs by ethnic group. Possible scenarios for future ethnic fertility are then identified, based on the evidence of previously estimated trends.

The authors conclude that it is possible to devise methods for projecting fertility within the United Kingdom, using existing data sources. In particular, they consider that the Longitudinal Study – with 2001 Census data linked in – is likely to provide, for the first time, a sufficient sample from which to estimate fertility by ethnic group and country of birth.

External migration

This seventh chapter, by John Salt, focuses on statistical information on international migration to and from the United Kingdom, with particular reference to the minority ethnic populations. Immigrants are often portrayed as the visible minority groups who settle in this country, which obscures the fact that immigration covers a wide range of migrant types and motives. Immigrants and ethnic minorities are very different conceptually, and the migration patterns of ethnic groups are quite distinct from those of all international migrants who arrive in, or leave, the United Kingdom. Nevertheless, it has been the immigration of minority ethnic populations, particularly from the New Commonwealth, which has largely driven immigration policy over the years. Migrants and migration of minority ethnic groups do share one feature; they are both poorly captured in official statistics, which are drawn from a variety of different sources, most of which are based on administrative systems concerned with immigration control, rather than the recording of migrants' characteristics.

The author then considers the available data sources that yield statistical information on international migration. The census is of limited use, since it does not ask about country of origin or nationality, even though place of residence one year ago is recorded. The availability of ethnic group since 1991 is not much help as increasing numbers of the minority ethnic populations have been born in the United Kingdom and are British citizens. The Labour Force Survey is, however, an important source for both stock and flow data on international migration. It is possible to analyse international migration by ethnicity, and also by year of entry into the United Kingdom, although sample sizes could well prove inadequate for some analyses.

The International Passenger Survey (IPS) is perhaps the most valuable source, since it is the only one providing data on both immigration and emigration. Besides recording the basic demographic characteristics of migrants, the IPS asks about citizenship, country of origin and destination region though the overall sample size of migrants at 2,500 is quite small. IPS estimates form an important element in analysing international migration, though they have to be adjusted to allow for the number of people who stay more than one year inside or outside the United Kingdom, contrary to their originally declared intention. The IPS does not ask about ethnic group, but some estimates can be inferred from grouped countries of origin and destination.

Home Office statistics on immigration control and settlement are a by-product of immigration control, with acceptances for settlement essentially measuring longer-term immigration. Not surprisingly, the statistics vary from year to year depending on a variety of factors, not least of which are changes in Immigration Rules.

The author then considers the adjustments that have been needed to migration estimates, and the reader unfamiliar with the subject will be struck by the fact that the different data sources have, to some extent, both to act as cross-checks one with another and also to complement one another. In addition, very pragmatic and often subjective adjustments have to be made to take account of some very practical features of actual immigration behaviour and its measurement. A review of migration trends since the early 1980s is then given, distinguishing inflows and outflows, and for British and non-British citizens entering the country.

The author concludes with a discussion of the influences that will shape the future pattern of immigration. Legislation, liberalisation, international patterns of labour and skills supply and demand will all undoubtedly play their part. Whether choice of marriage partners from ethnic origin countries will continue to the same extent, and whether short-term international migration for work will become an increasing tendency are hard to predict. Overall, the conclusion is that there will be considerable need (and leeway) for making subjective judgements on how the agreed total net inflow and outflow in the future migration assumption might be decomposed by ethnic group.

Internal migration

The eighth and penultimate chapter by Anne Scott and Lucy Vickers informs us on the available data on internal migration by ethnic group. As such, it is the logical sequence from the last paper, since while information on external migration by ethnic group is needed to produce national ethnic population projections, internal migration by ethnic group is needed to produce the corresponding subnational projections.

The authors explain that in the United Kingdom there is no compulsory notification of people moving within the country. It is therefore necessary to estimate internal migration using proxy and survey data; for small subpopulations such as ethnic minorities, the task is particularly difficult.

The authors then review the different data sources that yield information on internal migration. The census is the best data source since it includes a question that allows everyone whose address one year previously was different to be distinguished. The 1991 Census also asked everyone's ethnic group so that internal migration for 1990–91 – and, subsequently, for 2000–01 – can be measured at the smallest scale.

Patients changing General Practitioners when they move is used as a proxy measure of internal migration. Estimates of internal migration by the former Family Health Service Authority areas are produced, and further work has enabled corresponding estimates to be produced for all Local and Health Authority Areas. Currently, ethnicity is not recorded on the patient registers from which these estimates are derived.

In general, the sample numbers in social surveys are too small to yield reliable estimates of internal migration, although the Labour Force Survey is the largest of all and can provide some information on the internal migration patterns of the minority ethnic populations, even though the sampling variability can be large.

The authors also outline how combining data sources can assist in formulating migration assumptions by ethnic group. For example, in the Bradford ethnic population projection, both data from patients changing doctors and a names analysis to distinguish different ethnic groups enabled the migration assumption to be based on extra background information. Another possibility is that for each patient who is a Longitudinal Study member, the patient's ethnic group could be obtained from the Longitudinal Study using the patient's NHS number to link to the records.

Overall, the authors conclude that when both 1991 and 2001 Census migration results are available, there will be sufficient information to make population projections by ethnic group for the regions and countries – and also for local authority districts and unitary authorities if broader categories of ethnic groups are used.

Mortality

The ninth and final chapter, by Seeromanie Harding and Rasaratnam Balarajan, considers the nearest available data to mortality by ethnic group and discusses some issues that require methodological caution when analysing the data that are available.

In general, there is little mortality data by ethnic group, and most studies on England and Wales data use country of birth as a proxy. The decennial analyses – using mortality data during a three- or five-year period centred around the census year – have been undertaken on this basis. However, with the passage of time, the proportion of second and later generation immigrants has grown – and so has been omitted from the analysis.

The Longitudinal Study, (LS), a record linkage study which started with a 1 per cent sample of the population in England and Wales in 1971, enables the mortality of migrant groups to be explored, although sample numbers can be small. The 1971 Census asked everyone not only their own country of birth, but also that of both their parents, so that the LS can, in theory, permit *generational* comparisons in mortality, though sample numbers make such an analysis possible only for certain countries of birth. These countries are, for example, Ireland and Scotland, for which the numbers of migrants over the years have been very large, and which therefore ensure that the 1 per cent LS sample numbers are sufficiently large to give reliable results. However, the LS data set will increasingly be able to yield findings on differential mortality since it is now 30 years since the first adults – and migrants – were linked into the study.

The authors then give a very useful review of the possible pitfalls in interpreting estimated mortality rates and standardised mortality ratios. For these derived from combining mortality from death registration with census data – that is where the records are *not* linked – problems have been identified of ethnic misclassification: for example the misclassification at death of older people who were recorded as having been born in India but who were born in what is now Pakistan. Also, Caribbean-born people are usually assumed to be of African origin, but a small, though significant, proportion is of Indian ancestry.

Other areas of potential misinterpretation concern those born in the Indian subcontinent who were the children of British expatriates and return migration biasing mortality rates. Other issues addressed are generational differences in mortality and, related to that subject, changes in mortality with increased duration of residence in the United Kingdom.

Overall, it is clear that most of the work performed on mortality differentials has been for epidemiological purposes rather than for 'demographic account' purposes – and the availability of the necessary information has severely restricted what can be analysed and concluded. The formulation of mortality assumptions by ethnic group will therefore at best be based upon very crude evidence.

General conclusions drawn from the chapters

Overall, the authors express a consistent opinion that producing projections by ethnic group *is* a feasible proposition – and that the cohort component model – or one of its developed forms – is the most appropriate one to employ.

Mike Murphy has documented the advantages. In addition, the authors generally conclude that there *are* sufficient data from which to formulate the three main sets of assumptions, although it is evident that information on past trends is particularly sparse. The challenge is all the greater because the data input requirements to the multi-regional projection model are large, so that there is a considerable mismatch between the volume of data needed and that available. However, as Phil Rees has indicated in his chapter, aggregated data can be used in such circumstances, although, obviously, even greater reliance would be made on subjective judgement.

These fundamental considerations really point to the fact that making projections by ethnic group is not simply a case of it being feasible or infeasible, but being able to make them with an acceptable or an unacceptable degree of uncertainty. Even this conclusion is somewhat disingenuous, since one cannot assess the extent of the associated uncertainty without the information that would confirm whether they were feasible or not.

Given that there is clearly a need for ethnic projections, the solution appears to be much more modest than is the case with traditional projections with regard to the details and the time span of the projections – at least for the *published* results. There is scope for expert evaluation of the level of detail – and how far the projections might be produced into the future – *that would be permitted by the assumptions that would have to be formulated*; that is, the uncertainty surrounding the assumptions would determine the practical characteristics of the projections. Ludi Simpson has made suggestions for such practicalities.

Two other reflections should also be considered: from Marian Storkey's review it is clear that a number of other countries have produced corresponding projections with data scarcely better than that existing in the United Kingdom; and that, from the work of all the chapter authors, it is clear that projections by ethnic group in the United Kingdom would be placed on as good a theoretical and practical basis as possible – all of which has been well documented in this volume.

This is a blank page

Projections of the population by ethnic group: a sufficiently interesting or a definitely necessary exercise to undertake?

John Haskey

Summary

This paper considers the range of needs that would be met by having available projections of the population by ethnic group, and, in particular, the role these projections would usefully play in informing current and future policy issues. The paper first identifies and lists the main uses, and potential users, of such projections and gives an appraisal of the likely applications in each context. Some of the challenges involved in using the main data sources – the census and surveys – for making such projections are then addressed; the paper considers what is measured by asking questions on ethnicity, and discusses the implications of ethnicity being self-assessed – two important considerations that affect the potential usefulness of ethnic projections.

Introduction

Why population projections by ethnic group?

The 'traditional' population projections – those by age, sex and perhaps especially those by area – have been used for many years for a variety of planning purposes, particularly in the fields of housing, health and education, employment and transport. In this paper, the uses of population projections by age, sex, and ethnic group will be considered. An initial and very pertinent question is therefore whether – and how – these projections would usefully supplement the 'traditional' projections – whose uses have long been accepted as vital.

In essence, the most compelling reason centres on identifying those sections of the population towards which current policy issues are particularly directed – and certainly there has been increasing evidence that the minority ethnic populations experience disproportionate social exclusion, which, in practical terms, means poorer than average housing, health, education, and employment. Measures and policies are being put into place to achieve and maintain equality in these fields both now and in the future. On this basis, projections by ethnic group would therefore appear to be a most useful supplement to the existing set of population projections. The paper explores this subject and describes the range of likely uses of projections by ethnic group.

Uses of estimates and projections of the minority ethnic populations

In many instances, the uses of ethnic population *projections* would mirror those of the corresponding population *estimates*. For example, an important use of these population estimates is to monitor racial discrimination and measure equality of opportunity, and the corresponding projections may be used in similar vein – to estimate the likely future proportions of the different minority ethnic populations in the overall population to determine the monitoring targets likely to be required in the future. Such scenarios should also help the understanding of how present and future targets relate – and the fact that there need not necessarily be a steady transition from one to the other. Either way, any form of change may necessitate appropriate policies to be formulated, and there is considerable advantage in being aware in advance of likely trends.

Projections are also needed in their own right for a number of other purposes – that is, ones which are not just extensions of those requiring population estimates. These mostly involve future *numbers*, rather than *proportions* they will form of the entire population. As an example, many of the minority ethnic populations are currently younger than the White population, and the future numbers with given health conditions known to be disproportionately suffered by particular minority ethnic populations in old age can be estimated – as can the future numbers of deaths from specific causes amongst the different minority ethnic populations. Such analyses are needed in planning medical care facilities and the number of treatment centres. Other examples include long-term planning needs at national level – which range from resource allocation exercises to the provision of a variety of services – in which knowledge of the likely future numbers of the population by ethnic group would assist each planning process achieve its goals. Before considering further examples, however, a review will be given of the broader context of the need for all kinds of statistical information on the minority ethnic populations – in order to obtain a more complete picture and a better appreciation of the potential role of projections.

Current policy on minority ethnic issues and implications for improved statistical data

There have been a number of recent policy-driven initiatives relating to race in Government Departments. The Home Office recently published a report, *Race equality in public services* (Home Office, 2000):

'The promotion of race equality within the provision of services to the public is a key priority for this Government. In support of this priority, the Government is engaged in a programme of initiatives ranging from legislative changes to the Race Relations Act to the introduction of a new race equality grant. The publication of the performance data within this document will provide a quantifiable way of demonstrating our progress and will help to make race equality a core issue for Government. It will allow us to bring together the key issues for ethnic minority communities in one document.'

In order to meet the Government's objective of promoting race equality, a number of key areas of public sector service provision were focussed upon, including those of: economic activity; education; health; law and order; housing; and local government.

There is also a considerable demand for improved population *estimates* by ethnic group, the most important being for those by local area (Haskey, 1991). The recent ONS *Population statistics review* (ONS, 1999) has acknowledged this need, and plans to take advantage of the proposed increase in the Labour Force Survey sample size to provide these more detailed estimates which are needed for a variety of planning and analytic purposes.

More generally, the Government has recognised the crucial need for better information on many

issues to assess and address the problems existing in deprived neighbourhoods. In a report (Cabinet Office, 2000a) on obtaining better information for neighbourhood renewal, the Government has stated:

'Comprehensive and up to date information about deprived neighbourhoods is crucial to the National Strategy for Neighbourhood Renewal … Information is one of the key tools in the drive towards a more inclusive society …'

One of the main sets of recommendations concerns ethnicity – the desirability of obtaining more frequent and comprehensive data by ethnicity across a range of subjects. One specific recommendation, for example, is that, following the 2001 Census, a time-series of small area data should be provided, which is capable of analysis by ethnicity, amongst other factors.

The general need for *projections* was well argued in an earlier paper (Patel, 1999) by the Social Exclusion Unit who co-ordinated the production of the report:

'Prevention and prediction

Good diagnosis relies on the ability to see and predict trends … Government will be unable to keep up with and address rapid changes … without some means of being able to see what is coming up ahead. Poor quality data has led policy makers at a national, local and neighbourhood level to be reactive in their response to social exclusion, rather than put in place preventative strategies. Lack of data leads to government always playing catch up – dealing with problems after they have occurred … A solid information base is needed for predicting future trends and taking preventative action.'

The report on better information for neighbourhood renewal was prepared by a 'Policy Action Team' and co-ordinated by the Social Exclusion Unit. In all, there were 18 different Policy Action Teams set up to report on different aspects of social exclusion, and each was asked to cover their subject for people of different ethnic groups. As a result, a considerable body of evidence has been accumulated on the disproportionate exclusion experienced by the minority ethnic populations, and the actions currently being taken to remedy it. This evidence, together with the related recommendations, which include those on improvements in data on ethnicity, has recently been published as a further report on minority ethnic social exclusion issues (Cabinet Office, 2000b).

Through these Government initiatives, which aim to promote and uphold racial equality, one particular recommendation has been made to add ethnic group to existing administrative and other data on a variety of subjects. Using such data in conjunction with population estimates by ethnic group is potentially an important and developing field. With additional information on projections, the likely future numbers of the minority ethnic population in particular situations, or experiencing certain events, could be calculated, based on the assumption of current prevalence levels, or current rates, continuing unchanged. Such simple analyses could supply answers to pertinent 'what if?' types of questions.

Quite apart from these possibilities there is a wide range of other uses of projections by ethnic group for policy and planning purposes and research on:

- medical, health, epidemiological and related issues;
- language and culture;
- service provision;
- education;
- labour market issues;
- finance and economic matters;
- family, inter-generational, and old age issues;
- travel, migration and immigration.

Besides the projection results being available for general use, other users are likely to be as follows:

- Parliament.
- Central Government.
- Local Government.
- The NHS, NHS Health Trusts, and private medical insurance.
- Personal social service providers – to ensure equal access to, and services from, these providers.

- Public bodies, such as the Commission for Racial Equality.
- Trade unions, and employers' organisations – to monitor patterns of employment and unemployment for the different ethnic groups.
- Business and commerce – to ensure that recruitment, training and retention policies are fair and transparent, and that educational skills and qualifications are matched to appropriate jobs.
- Marketing of consumer goods and services – to provide products and services that reflect the increased diversity of cultural norms.
- Insurance and finance companies.
- Academic researchers.
- Educational organisations and schools – to reverse the growing diversity of educational achievement between children of different ethnic groups.
- Students.
- Voluntary bodies and agencies.
- Charities.
- Religious, language and cultural organisations – to recognise that language training and translation facilities are needed to avoid exclusion on linguistic grounds.
- The media.

Further details and discussion of the likely range of uses of projections of the minority ethnic populations

The above lists give a summary, in briefest form, of the potential uses, and users, of the projections and certainly some of them deserve fuller attention, if only to identify specific issues. Before doing so, it is appropriate first to consider the wider context and developments that are likely to take place in society. In a recent paper, Penn (2000) has provided a summary of the likely changes in society to which these projections will relate. The paper examines aspects of demographic change – including those due to ethnicity, migration and fertility – and makes conjectures on their wider impact on society, concluding that issues of ethnicity will become of increasing importance.

His thesis, with only minimal changes to his wording, is as follows:

'Britain is already a complex multi-ethnic society… which is likely to increase in complexity and scale over the next twenty-five years ... Linguistic and ethnic diversity will increase as a result of a number of factors (*most notably the effects of international migration and persistent differentials in ethnic fertility*). It is also unlikely that there will be a diminution in the number of asylum-seekers and refugees wishing to enter Europe – and by extension, Britain … This will raise a whole series of difficult issues in the spheres of education, training, and employment, as well as contributing to the development of new cultural forms and practices.'

As noted above, the key issues are those of racial inequality and social deprivation and exclusion – which are suffered disproportionately by each of the minority ethnic populations. Projections are decidedly necessary to provide a scenario for the future, define national targets, and track developments towards meeting them. Only with the possession of projected numbers can actions be formulated, and progress monitored adequately in all the different spheres where situations need improving or likely future developments need clarifying and quantifying.

Some particular issues, which are likely to be of growing significance in the future, are as follows:

- Appropriate numerical representation in various public services, notably the Police, the Judiciary and the Armed Forces.
- Disparity of individuals', families' and households' incomes.
- Reducing excess mortality due to certain causes amongst specific minority ethnic populations.
- Making allowance in the provision of medical services for variations in the incidence and prevalence of morbidity, including mental health, and formulating policies to reduce above-average morbidity.
- Differential incidence of school exclusions.

- Addressing the issue of separate schools for different religious groups.

If projections of the population by ethnic group are to be made, a number of considerations need to be addressed. The first is whether projections can indeed be made with existing data, and if judged possible, whether other difficulties can be overcome. These issues will now be considered.

Can projections be made with existing ethnic group data?

In order to make projections of the population by ethnic group, four main data sets are required, at least if the standard component method of projection is used:

1. a base population by ethnic group;

past trends in:

2. fertility by ethnic group;

3. mortality by ethnic group;

4. migration by ethnic group.

(These past trends, in the form of rates, are needed to formulate sets of assumptions of corresponding *future* rates.)

The One Number Census would be the main source for the base population. In intercensal periods, population estimates are made (Schuman, 1999) from cross-sectional sample surveys, in particular the Labour Force Survey (LFS), which asks the same ethnic group question as the census.

In addition, the Census would most likely play an important part in estimating past recent trends in the three components (fertility, mortality, and migration) via analyses of the ONS Longitudinal Study (LS), which incorporates data from the past three censuses including the 1991 Census, the first to include a question on ethnic group.

Hence, for much of the statistical data needed to make projections, the ethnic group classifications used in the 1991 Census – and also in the 2001 Census – really determine, not only the different ethnic groups which are distinguished, but also how the information on them is collected.

It should be recognised, however, that the ethnic group question in the census is intended to serve many purposes simultaneously; in particular, the question is **not** formulated specifically to provide a detailed classification of ethnicity for demographic or research purposes. As part of its broad remit, the question must be easy to ask and practical, not only in the census but also in other situations, and so capable of providing a suitable standard.

The ethnic group question

In both the 1991 Census and the LFS, respondents are invited to select from a list of nine categories, the ethnic group to which they consider they belong – see the Appendix. There therefore appear to be two factors determining the ethnic group which is recorded for each respondent: their own choice of how they view their own ethnicity, and the list of options which are presented to them. The first consideration is the most straightforward; the ethnic group that each person chooses as his or her own is intrinsically the ethnic group of self-identity, rather than being ascribed by anyone else. The second consideration is apparently not so clear-cut, in that the ethnic group options presented to the respondent are *not completely* ones of self-identity in the sense that the respondent has no say in the names or the number of the different alternative ethnic groups in the 'menu'. Therefore, the freedom the respondent has to select their own ethnic group is constrained and influenced by the options on offer.

In reality, whether the data-collecting occasion is that of the census (Teague, 1993) or of the LFS, the ethnic groups presented are the result of much piloting and pre-testing before a standard list is put into practice. Extensive field work was carried out before the 1991 Census (Sillitoe and White, 1992; Sillitoe, 1987a and 1987b) and the 1991 Census ethnic classification was subsequently adopted by the LFS as part of the harmonisation of the ethnic group question. A large part of the piloting centres on testing alternative wording of the question and the named categories to accord with people's own preferred ethnic descriptions of themselves. The underlying aim in drawing up a list of ethnic group options is therefore consistent with the self-identity aspect implied by respondents choosing their own ethnic group. In practice, the list of apparent *answers* is effectively a list of *questions*, with the respondent successively asking

himself/herself 'Am I White, Black-Caribbean …?' Answering 'Yes' to one particular ethnic group involves identifying with that ethnic group.

Perceptions of ethnic group

However, it should be recognised that individuals' and groups' own perceptions of their ethnicity are also influenced by the views of other minority ethnic groups – as well as of the majority White population – towards them. Another important factor is how the White population perceives the variety in distinguishable ethnic groups, being influenced by a number of characteristics (Haskey, 1997) including both physical and non-physical ones, such as skin colour, culture and language. This factor probably accounts for the fact that only three ethnic groups are identified amongst those from the Indian subcontinent (Indians, Pakistanis, and Bangladeshis) – rather than separately identifying, say, Gujaratis, Sylhetis, Punjabis, and Azad Kashmiris – although part of the explanation may lie in not wishing to have too numerous a classification. In reality, there is a dynamic process of interchange between the self-definition of the minority groups and the perceptions of society as a whole.

Implicit from the intrinsic nature of ethnicity is the fact that the ethnic groups, however defined or measured, will tend to change over time – so that, quite legitimately, for a proportion of the population, a person may record themselves as one ethnic group at one time and another on a subsequent occasion. Such changes depend upon social and political attitudes and developments; for example, 'Black' was an unacceptable term at one time, but is now one that is embraced by the individuals in the groups concerned. This, and earlier examples indicate that:

an ethnic group classification is only valid for the period and context in which it is used.

More generally, when the previous set of ethnic population projections was made in 1976 (Office of Population Censuses and Surveys, 1979; Haskey, 2000), the term for minority ethnic population was the population of 'New Commonwealth and Pakistan ethnic origin', but the terminology of ethnicity had changed by the time the final year of the projection arrived – which was 1991.

In the piloting of alternative names for ethnic groups to include in the list of options in the census, some inconsistencies or difficulties inevitably arise. The 1991 Census did not include a specific 'mixed' ethnic group category. Up to the mid-1980s, the various field trials had shown that people of mixed descent often preferred *not* to be distinguished as a separate group; instead they usually identified with the ethnic group of one of their parents – usually the father (Sillitoe, 1987a). As a result, an attempt to classify all persons of mixed descent in the same way was abandoned, and a guidance note was added to the 1991 Census question: 'if the person is descended from more than one ethnic or racial group, please tick the group to which the person considers he/she belongs, or tick the "Any other ethnic group" box and describe the person's ancestry in the space provided.'

Evidence from the LFS suggested, however, that, by 1991, the population of mixed ethnic origin was increasing with an annual growth rate (Haskey, 1992) above the minority ethnic average. Probably as a result of trying to estimate the numbers by indirect means, this group could not be fully accounted for in the 1991 Census results. More recently, fieldwork to determine a revised ethnic group question for 2001 showed that a 'mixed' category was acceptable, provided that an opportunity was given to record the relevant details, as a written description (Phillips, 1999).

This finding illustrates that those of mixed ethnicity may well respond differently at different times – depending upon their changing views of their own ethnicity, which in turn will be influenced by society's changing perceptions – and ability to distinguish such ethnicity. In fact, far from being a single ethnic group, those of mixed ethnic origin consist of a multitude of different ethnic backgrounds. In recognition of this diversity, there was not only a separate mixed ethnic group question in the 2001 Census, but three of the most important kinds of combinations of origins were distinguished within it (Her Majesty's Government, 1999; Teague, 2000).

Another example is the tendency of some Black-Caribbeans to describe themselves as Black British (Peach, 1996; Owen, 1996) to emphasise, if not for themselves then for their children, that they were born in Britain, belong in Britain, and are therefore British. The desire to use the term might well also signify the assertion that being 'Black' and 'British' are not inconsistent. The major difficulty with the ethnic group 'Black British' is that it conceals the further details of the

person's origins, preventing the finer classification of ethnicity which most users require. A similar development has occurred amongst younger members of the Indian, Pakistani and Bangladeshi communities who are beginning to prefer the term 'British Asian'.

However, it can be argued that aiming for greater detail than is given by 'Black British' or 'British Asian' is interfering with the concept of unfettered self-classification, even though 'British' is a problematic and ambiguous term – which can refer to nationality, citizenship, and, for many, ethnicity (McCrone and Kiely, 2000). In fact, this problem has apparently been largely overcome in the 2001 Census (ONS, General Register Office for Scotland and Northern Ireland Statistics and Research Agency, 1999) by what is essentially a two-stage system of choosing ethnic group from the census schedule classification; one of the first stage choices includes the category: Black or Black British, whilst the second stage refines that choice into several options, including 'Black-Caribbean'.

Nevertheless, this device may prove only a short-term solution – which allows some consistency and continuity with the 1991 Census classification – if a growing proportion essentially do not wish to elaborate on 'Black British' or 'British Asian', or do not identify with 'Black-Caribbean'. More importantly, the basis or implied basis, of the ethnic group classification can influence the pattern of response. The implications for projections that start from a base population are obvious.

It should also be mentioned that there is a facility for respondents to specify their ethnic group should they not be satisfied with any of the suggested descriptions in the ethnic group classification. In fact, in 1991 there were two such opportunities for writing in one's own terminology of one's ethnic group: firstly, if one were 'Black', and none of the 'Black' ethnic groups was deemed a satisfactory description, and secondly if none of the entire set of ethnic group descriptions was judged appropriate. A similar pattern – and ability to specify one's own ethnic group if so wished – was also incorporated in the 2001 Census. (In fact, there were five, rather than two, separate opportunities (Her Majesty's Government, 1999; Teague, 2000) to write in one's own description – under, in turn: any other White/Mixed/Asian/Black/Other ethnic group.)

Can projections be made by ethnic group using available data?

In making projections by ethnic group, an important question is whether ethnic group as collected in the census and LFS provides objective or subjective information. If it is the former, then projections by ethnic group should be straightforward – or at least as straightforward as those by, say, legal marital status (even if extra attention needs to be paid to children of parents of different ethnic groups). If, however, it is the latter, projections may range from mildly difficult to impossible, depending upon the degree of subjectivity, and the rapidity with which that subjectivity results in the categories changing over time.

At first sight, ethnic group being self-ascribed, and based on self-identity, appears to be entirely subjective. However, the named ethnic group descriptions provided, that is, the precoded answers – or *precoded questions* – consist of a combination on colour and geographical region of ancestry – both of which are reasonably objective for the majority of the population. Furthermore, rather more pragmatically, respondents to the 1991 Census answered by and large in a consistent way, and judging by the profile of answers to the LFS since 1992, there have been no discernible shifts from one ethnic group to another. Nor has there been strong evidence of a rejection of any of the ethnic group terms. Indeed, the ethnic groups that appeared in the 2001 Census – the result of extensive field-testing – show strong continuity with those appearing in the 1991 Census.

Consequently, the basis on which ethnic group is collected is probably sufficiently objective to contemplate making projections. The strengthening of the 2001 Census to overcome the difficulties experienced in 1991 also augurs well for the feasibility of a projections exercise.

In fact, there were a number of other proposals for, and discussions of, the best ethnic group question for the 2001 Census (Ballard, 1996 and 1998; Berthoud, 1998; Aspinall, 2000). It was decided (ONS, General Register Office for Scotland and Northern Ireland Statistics and Research Agency, 1999) after a question testing programme, that there should be a balance between continuity with the 1991 classification and making the desirable improvements identified as necessary in the process of analysing the 1991 Census – most particularly the detailed

analyses contained in the four volumes *Ethnicity in the 1991 Census* (Office of Population Censuses and Surveys, 1996; ONS, 1996a, 1996b and 1997).

What are the implications for projections?

The above considerations have been discussed in some detail since they have obvious implications for the practicability, robustness and interpretation of minority ethnic population projections results. Their relevance is all the more important if the basis, or definition, of ethnic group is changed. While there are a number of uncertainties attached to 'traditional' population projections – those by age and sex – there are additional uncertainties involved with **ethnic** projections:

(a) Uncertainties resulting from the imprecision of ethnicity

- Changing perceptions of ethnic identity, including the distinguishing of new ethnic groups – including those within the White population – and the possible merging of some 'traditional' ethnic groups.
- The growing trend for different groups within the population of the United Kingdom to wish to have their identity recognised. Devolution has probably added extra impetus to this trend.
- A shift in emphasis in the definition of ethnic group, from one predominantly based on colour and country of ancestral origin, to one placing greater weight on language, culture, and religion.
- The greater variation in ancestry in those of mixed ethnic origin – resulting in a very heterogeneous group in demographic and other characteristics.
- The greater likelihood of those of mixed ethnic origin answering differently on different occasions, depending on the context of the situation.
- The added complication of births of children to parents of different ethnic groups being considered by their parents as belonging to one or other of their ethnic groups – or indeed to a third.
- The fact that when they grow up, children may not consider they belong to the same ethnic group which their parents had earlier ascribed to them.

(b) Uncertainties connected with demographic considerations

- Greater difficulties in establishing past trends in fertility, migration and mortality, from which to formulate assumptions on future corresponding vital rates.
- Greater sensitivity of projection results to migration assumptions since issues such as asylum seekers, refugees, 'economic migrants' and return migration are especially dependent upon ethnic group – and immigration policy and controls may also have a differential impact.
- Uncertainty whether the fertility of the different ethnic groups – or only some – will converge with that of the majority White population, and, if so, at what pace.
- How the socio-economic situation of the different ethnic groups will influence their future patterns of fertility and mortality.
- Because the minority ethnic population has a younger age structure than the White population, the cumulative effect of relatively small variations in fertility assumptions for the minority ethnic population is likely to be relatively large.
- The need to reconcile the results of the 'traditional' population projections – those by age and sex – with those derived separately for each ethnic group.
- The fact that projection results are the implications of the chosen set of assumptions, rather than a definite prediction, which, if so viewed, may become self-fulfilling.

These are the main issues that need addressing and resolving. As mentioned above, a previous, relatively straightforward, projection exercise yielded acceptable results – though it should be added that the projections did not attempt to distinguish projected populations *by ethnic group*. Certainly, two previous parallel exercises both to estimate annual growth rates, one of the minority ethnic population as a whole (Shaw, 1988), and the other of each of the different minority ethnic populations (Haskey, 1992), showed, not surprisingly, that the latter set was more difficult

to estimate in practice, primarily because of the greater number of factors to take into account, and also because each of those factors had to be quantified or estimated on sparser quantities of information.

Conclusions

From the above considerations, it is concluded that, for a wide range of purposes, projections of the minority ethnic population are a very necessary element in the armoury of available national demographic data. Just as the traditional kind of population projections – those by age and sex – have long been accepted as indispensable for a variety of policy, planning and research purposes, so too should those by ethnicity. Furthermore, there is some qualitative evidence that ethnic group as collected in the census and the LFS is sufficiently objective to contemplate making projections.

References

Aspinall P J (2000) The new 2001 Census question set on cultural characteristics: is it useful for the monitoring of the health status of people from ethnic groups in Britain? *Ethnicity and Health*, vol 5 no 1, pp 33–40.

Ballard R (1996) Negotiating race and ethnicity: exploring the implications of the 1991 Census. *Patterns of Prejudice*, vol 30 no 3, pp 3–33.

Ballard R (1998) Asking ethnic questions: some hows, whys and wherefores. *Patterns of Prejudice*, vol 32 no 2, pp 18–37.

Berthoud R (1998) Defining ethnic groups: origin or identity? *Patterns of Prejudice*, vol 32 no 2, pp 53–63.

Cabinet Office, Social Exclusion Unit (2000a) *National strategy for neighbourhood renewal: report of Policy Action Team 18 – better information*, TSO: London.

Cabinet Office, Social Exclusion Unit (2000b) *Minority ethnic issues in social exclusion and neighbourhood renewal: a guide to the work of the Social Exclusion Unit and the Policy Action Teams so far*, Cabinet Office: London.

Haskey J (1991) The ethnic minority populations resident in private households – estimates by county and metropolitan district of England and Wales. *Population Trends* 63, pp 22–35.

Haskey J (1992) Demographic characteristics (and annual growth in the size) of the ethnic minority populations of Great Britain, in Bittles A H and Roberts D F (eds), *Minority populations: genetics, demography and health*, Studies in biology, economy and society, Macmillan: Basingstoke/Galton Institute: London.

Haskey J (1997) Population Review: (8) the ethnic minority and overseas-born populations of Great Britain. *Population Trends* 88, pp 13–30.

Haskey J (2000) Demographic issues in 1975 and 2000. *Population Trends* 100, pp 20–31.

Her Majesty's Government (1999) *The 2001 Census of Population*, White Paper (CM 4253), TSO: London.

Home Office (2000) *Race equality in public services – driving up standards and accounting for progress*, Home Office: London.

McCrone D and Kiely R (2000) Nationality and citizenship. *Sociology*, vol. 34 no 1, pp 19–34.

Office for National Statistics (1996) *Ethnicity in the 1991 Census, volume 2: the ethnic minority populations of Great Britain*, Peach C (ed), HMSO: London.

Office for National Statistics (1996) *Ethnicity in the 1991 Census, volume 3: social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration*, Ratcliffe P (ed), HMSO: London.

Office for National Statistics (1997) *Ethnicity in the 1991 Census, volume 4: employment, education and housing among the ethnic minority populations of Britain*, Karn V (ed), TSO: London.

Office for National Statistics (1999) *Population statistics review*, ONS, Population and Vital Statistics Division: Titchfield.

Office for National Statistics, General Register Office for Scotland and Northern Ireland Statistics and Research Agency (1999) *2001 Census output: proposed data classifications*, ONS: Titchfield.

Office of Population Censuses and Surveys (1979) Population of New Commonwealth and Pakistani ethnic origin: new projections. *Population Trends* 16, pp 22–27.

Office of Population Censuses and Surveys (1996) *Ethnicity in the 1991 Census, volume 1: demographic characteristics of the ethnic minority populations*, Coleman D and Salt J (eds), HMSO: London.

Owen D (1996) Black-Other: the melting pot, in Peach C (ed), *Ethnicity in the 1991 Census, volume 2: the ethnic minority populations of Great Britain*, ONS, HMSO: London, pp 66–94.

Patel A (1999) *The work of the Social Exclusion Unit on better information.* Paper presented at the Statistics Users' Council/CEIES Joint Conference on Social Exclusion Statistics held at the Royal Society, London, 22 November 1999.

Peach C (1996) Black-Caribbeans: class, gender and geography, in Peach C (ed), *Ethnicity in the 1991 Census, volume 2: the ethnic minority populations of Great Britain*, ONS, HMSO: London, pp 25–43.

Penn R (2000) British population and society in 2025: some conjectures. *Sociology*, vol 34 no 1, pp 5–18.

Phillips M (1999) *Ethnicity in the 2001 Census*. Talk given at the Black Employment Conference, 1999.

Schuman J (1999) The ethnic minority populations of Great Britain – latest estimates. *Population Trends* 96, pp 33–43.

Shaw C (1988) Components of growth in the ethnic minority population. *Population Trends* 52, pp 26–30.

Sillitoe K (1987a) *Developing questions on ethnicity and related topics for the Census*, OPCS Occasional Paper no 36, Office of Population Censuses and Surveys: London.

Sillitoe K (1987b) Questions on race, ethnicity and related topics for the Census. *Population Trends* 49, pp 5–11.

Sillitoe K and White P H (1992) Ethnic group and the British Census: the search for a question. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, vol 155 no 1, pp 141–63.

Teague A (1993) Ethnic group: first results from the 1991 Census. *Population Trends* 72, pp 12–17.

Teague A (2000) New methodologies for the 2001 Census in England and Wales. *International Journal of Social Research Methodology: Theory and Practice*, vol 3 no 3, pp 245–55.

Further reading

Additional issues relevant to those discussed in this paper are addressed in:

Modood T, Berthoud R et al (1997) *Ethnic minorities in Britain: diversity and disadvantage*, Policy Studies Institute: London.

(This volume provides comprehensive analyses, derived from the Fourth National Survey of Ethnic Minorities, on the position of the minority ethnic communities in British society. It reports on changes in such key features as family and household structures, education, qualifications and language, employment patterns, income and standard of living, neighbourhoods and housing. It also analyses some new topics, including health and health services, and cultural identity. Of particular interest, the volume addresses the issue of obtaining better information on ethnicity.)

Lieberson S and Waters M C (1988) *From many strands: ethnic and racial groups in contemporary America*, Russell Sage Foundation (for the National Committee for Research on the 1980 Census): New York.

(This volume describes the radical changes in the measurement of ethnicity in the 1980 US Census by collecting information on ancestry for all respondents, regardless of how long ago their forebears migrated to America, and by allowing respondents of mixed background to list more than one ancestry. While some findings lent support to the 'melting pot' theory of assimilation, other findings suggested the persistence of pluralism.)

Acknowledgements

Thanks are due to the referees, Dr Roger Ballard, and particularly to Professor Richard Berthoud who provided many helpful comments on an earlier draft. The opinions expressed, and any errors of fact or interpretation remain, however, the responsibility of the author alone.

Appendix: The ONS ethnic group classification

Since 1991, the ethnic group classification used in the 1991 Census has become the standard:

White;

Black-Caribbean;

Black-African;

Black-Other;

Indian;

Pakistani;

Bangladeshi;

Chinese;

Other.

In the 1991 Census, respondents who ticked either Black-Other or Other were asked to write in further details of their ethnic group, which, depending upon what they supplied, was either classified back to one of two new groups:

(a) Other Groups-Asian; or,

(b) Other Groups-Other (non-Asian).

In the LFS, respondents who respond to the interviewer's list of alternatives by giving either Black-Other or Other are asked to give more details; depending on their reply, their answer is either classified back to one of the main ethnic groups or else assigned to one of several new groups:

Black-Other (non-mixed);

Black-Mixed;

Other-Asian (non-mixed);

Other-Other (non-mixed);

Other-Mixed.

(Please see *The 2001 Census of Population* White Paper (Her Majesty's Government, 1999) for the 2001 Census ethnic group question.)

Chapter 1

A review of literature on ethnic group projections

Marian Storkey

1.1 Introduction

Ethnic group projections constitute a relatively new area of demographic research and consequently the literature that has been published is very limited. Much of it is very weak on detail of methodology and the assumptions used. However, this work is still able to inform the modelling process. This literature review provides a summary of work that has been undertaken both nationally and internationally on this topic.

In more detail the aims of the review are as follows.

- Firstly to summarise historical national population estimates which contribute to population projections by providing accurate base populations (Section 1.2). Trends in the estimated sizes of the minority ethnic populations highlight differential growth rates in the past – which provide information on likely future growth rates.

- Secondly past projections by ethnic group undertaken by OPCS (Office of Population Censuses and Surveys) are outlined in Section 1.3. These straightforward projections have not been updated. Their lack of sophistication – and the reluctance of OPCS to replace them with a more recent set – highlights the difficulties inherent in producing ethnic group projections.

- The handful of more recent population projections by ethnic group prepared by local government officers and academics are then considered in Section 1.4. Strengths and weaknesses of these models are discussed.

- Finally international work in this area is summarised in Section 1.5. Most of the countries included have used projection models with very simple assumptions about future patterns. The largest problems have been: the estimation of migration schedules; dealing with ethnic inter-marriage in terms of the identity of children; and dealing with changing ethnic identity over time. In some countries (but not in Britain) extra data are available from population registers that give continuously recorded information on residents, vital registration events by ethnic group, and more historical data to provide trend information.

All the work summarised here contain copious 'health warnings' about the projections – reminders of the difficulties of the task in hand.

1.2 Population estimates by ethnic group

Fundamental to any population projection is an accurate estimate of the base population. Population estimates for successive years enable growth rates to be estimated which can be used to predict future growth and also to evaluate previous projections.

There are a number of data sources that may be used to measure the size of ethnic minority populations in Britain. In evaluating these estimates, it is necessary to examine the methodology of their derivation and, if appropriate, the sample size of the data source because these are important determinants on the reliability of the estimates. The format of the ethnic group question and the categories used are also important. The main data sources in Britain are outlined below.

- Successive censuses (1841 to 2001) have measured the overseas-born population by asking for the respondent's country of birth. The 1971 Census also asked for the countries of birth of the respondent's parents. The 1991 and 2001 Censuses asked for ethnic group as well as country of birth.

- From 1969 the vital registration system recorded deaths by birthplace of the person, and births by birthplace of the mother.

- The General Household Survey (GHS) was the first large-scale survey to distinguish between 'White' and 'Coloured' respondents. In the early 1970s, the distinction was made by interviewers – it was not self defined. In 1971, the GHS identified 1,512 people as being 'Coloured' – which was equivalent to 2.3 per cent of the British population (Peach, 1982). This sample enabled only a very limited analysis of this group in the population.

- The National Dwelling and Housing Survey conducted in 1977–78 was the first national survey to collect direct information about the ethnicity of the population.

- The Labour Force Survey (LFS) provided the first set of regular and comprehensive information about ethnic minority populations in Britain. In 1979, 1981, and annually from 1983 to the present, it has asked respondents their ethnic group and also the ethnic group of other residents in their household.

The LFS can contribute to population projection by ethnic group in a number of ways. Base populations can be estimated using the LFS, and these can be validated against Census populations, which often form the main base population for projections (OPCS, 1986a). Also, intercensal growth rates can be estimated from the LFS, which can inform population projection growth trends (Owen, 1996a; Eversley et al, 2001). It must be noted that sampling error and lower response rates for certain ethnic groups need to be taken into account when using the LFS. A number of studies have assessed these effects (Owen, 1993; Sly, 1993).

Assessing ethnic origin from names has been undertaken since the 1970s. Names analysis was used on 1971 and 1974 births and deaths registrations as well as on a 1 per cent sample of those enumerated in the 1971 Census with certain New Commonwealth and Pakistan (NCWP) countries of birth (OPCS, 1975). The exact method used was not stated. This exercise was undertaken to relate birthplace to ethnic origin in order to exclude people born in the NCWP who were of White British ethnic origin, and to include people born in the UK who were of NCWP ethnic origin.

Revised estimates of the NCWP population for 1971 to 1974 and through to 1976 were published in 1977 (OPCS, 1977a). This publication contained more detail about the names analysis

undertaken. Over 6,000 census forms were examined and the ethnic origin of each person was assessed by examination of the respondent's surname or family name, birthplace and their parents' birthplaces. It was acknowledged that the method only worked for those who did not have 'British-style' names and so all those who were born, or whose parents were born, in the American New Commonwealth were assumed to be of NCWP ethnic origin. This study also referred to a periodic analysis of names that was undertaken on birth and death registrations to estimate ethnic origin proportions to apply to registration data.

Names analysis has been used in a number of studies and the accuracy of names analysis has been reviewed (Harding et al, 1999). The project used three data sets from Bradford and also Longitudinal Study members born in the Indian subcontinent. The Longitudinal Study is a record linkage study of census and vital event information for England and Wales. It covers approximately 1 per cent of all people born on four calendar dates during the year. The conclusions of the project were that analysis of names provided a better classification of South Asians than country of birth, but that there were limitations. The study reported that the computerised programme Nam Pehchan developed in Bradford produced good results, but more inaccuracies were evident when used against a national sample. These may have been caused by an incomplete name dictionary (which has since been updated), change of surname by women upon marriage, or the use of Anglicised names.

Estimates of ethnic minority populations can be updated using locally available sources of data. For example, Luton Borough Council has produced updated ethnic group estimates by using schools Form 7 ethnic data applied to mid-year estimate (MYE) population of school age children, ethnic figures from a skills audit survey applied to the MYE for working age people, and 1991 Census ethnic proportions applied to 1998 MYE figures for the pre-school age and retirement aged populations (Barton, 2000). The report acknowledges the possible inaccuracies of the updated figures because of the 'assumptions made' and 'the different data sources used and sampling errors'. Applying 1991 Census ethnic group proportions to MYE figures for the younger and older populations assumes no differential change by ethnicity over time in these age groups – which is unlikely. However, despite drawbacks in methodology, the figures are still likely to be a better reflection of the borough's ethnic composition than 1991 Census results.

Other estimation work has included estimating the size of 'second generation' populations, as in the case of the Irish population where no ethnic group data exists from the 1991 Census, only data for those people born in Ireland (recorded as Northern Ireland, the Republic of Ireland and Ireland – part not stated). Hickman and Walter (1997) estimated the Irish population from household data and derived a multiplier of around three, i.e. the overall size of the Irish population was around three times the size of those born in Ireland.

The work on estimates of different ethnic groups which is summarised here highlights a number of data sources which would be useful in the production of ethnic group projections. The 1991 and 2001 Censuses both included an ethnic group question, although different categories were used in the two censuses. These censuses provide important base data for projections by ethnic group because the almost complete coverage produces more accurate estimates of the minority ethnic populations than any survey. Growth rates in the minority ethnic populations between 1991 and 2001 calculated from census data will provide a useful guide to the differential rates of growth during that decade – and also give an indication of the expected future rate of growth, serving as a check for the plausibility of results from an ethnic group projections model.

Other nationally available data sets such as the Labour Force Survey provide useful data on intercensal change and can be used as a constraint for post-census ethnic group projections at regional and/or local area levels. Local survey data such as ethnic group of schoolchildren can be used locally as a check against projections in the same way that local data are often analysed by users to validate the total population projections produced by ONS.

1.3 OPCS projections

OPCS produced several sets of projections of the minority ethnic population in Great Britain from the early 1970s to the mid-1980s. No projections by ethnic group have ever been published by OPCS/ONS (Office for National Statistics) at subnational level. The national projections are examined in more detail below in order to note

data sources and methodology. They vary in complexity and the length of the projections period. This series of publications included back projections to 1966 and forward projections to 2001. Two simple estimates were published in 1970 (Hansard, 1970) and in 1977 (OPCS, 1977a). Four sets of projections were published in 1975 (OPCS, 1975), 1977 (OPCS, 1977b), 1979 (OPCS, 1979) and 1986 (OPCS, 1986).

The first estimate of future increases in the ethnic minority population was made in the late 1960s and published in Hansard in 1970. A Parliamentary Question was asked and in response an unsourced estimate was given that the population of NCWP ethnic origin in 1986 'was of the order of 2.5 million' (Hansard, 1970).

The first more detailed projections of the NCWP population by OPCS were published in 1975 (OPCS, 1975). They were short-term projections – updating 1971 Census based estimates to mid-1974, and back-projecting to 1966. They were not termed projections, but were instead labelled updated estimates although they are technically projections. The projections were calculated for the population of NCWP ethnic origin. The projections were based on the Census questions on birthplace and birthplace of parents.

Estimates of the numbers of births and deaths by ethnic origin were made by applying conversion factors based on an analysis of names. The factors converted births and deaths by birthplace to births and deaths by ethnic origin. A names analysis was completed using two data sources: registrations of births and deaths since 1969; births by birthplace of mother and father and deaths by birthplace of the deceased, and a 1 per cent sample of the 1971 Census of those born in certain NCWP countries. It was acknowledged that this procedure was subjective and a number of broad assumptions were made. For example it was assumed that all persons who had parents born in the West Indies were of West Indian ethnic origin.

Estimates of migration were calculated from the International Passenger Survey (IPS) based on migration patterns of citizens of the NCWP, including UK passport holders from East Africa. It was acknowledged that this definition did not equate exactly to NCWP ethnic origin and also that the sampling errors of the survey estimates were relatively large.

As births and deaths by birthplace were not recorded before April 1969, back projections were made by estimating the numbers of births and deaths by birthplace using 1969 data, and then making assumptions about the growth in the population between 1966 and 1969.

These projections provided updated population estimates for the early 1970s as well as new ones for 1967–70, and a revised estimate for 1966. The revised estimate of 1.016 million for 1966 replaced the one published in the Registrar General's Quarterly Return (OPCS, 1970) that used the 1966 Sample Census as the starting point for an estimate of 970,000.

In 1977 (OPCS, 1977a) a projection for the NCWP population in the mid-1980s was published. This rough projection suggested that the NCWP ethnic origin population would grow by 'about one half' from 1976 to 1986, to 2.66 million. Details of methodology were not included, although it was assumed that the growth rate of the NCWP population between 1976 to 1986 would be lower than the 5 per cent between 1971 and 1976.

The 1975 projections (OPCS, 1975) were later updated by projections published in 1977 (OPCS, 1977b) which used more detailed Census tables in a new set of projections for the period 1971 to 1976. These were essentially retrospective projections as they related to earlier years, but used assumptions for births, deaths and migration. These new projections suggested that the earlier ones were slightly overestimated. The estimate for 1971 was reduced by 130,000 to 1.360 million. The earlier projection for 1974 was 4 per cent higher than the later projection. Small changes in the assumptions were made in these 1977 projections. For example an 'arbitrary' 5,000 per year was added to the net migration of NCWP citizens, derived from the IPS in order to allow for possible under-recording due to 'language problems'.

A further set of projections was published by OPCS (1979). These detailed projections were for the period 1976–91, and a more simple projection was made to 2001. The projections were for the population of NCWP ethnic origin resident in Great Britain. A brief section was included on the projected size of the population of West Indian and Indian ethnic origin. The base population was the earlier mid-1976 estimate (OPCS, 1977b) with some adjustments. The mortality rates that were used were those of the population as a

whole, and these were applied throughout the projection period. It was decided that the number of deaths was small and that the projections would not be very sensitive to the choice of future mortality assumptions. Two fertility assumptions and two mortality assumptions were used.

The fertility calculations were fairly detailed. Three different subpopulations of women were considered: women born in NCWP living in Britain in 1976; women born in Britain of NCWP ethnic origin; and women born in NCWP who arrived in Britain after 1976. It should be noted that women of different ethnic groups within these subpopulations were all assigned the same fertility rates. The ratio of male to female births was assumed to be 104:100, the observed birth ratio for the NCWP population between 1971 and 1976.

Two fertility assumptions were used for NCWP-born women living in Britain in 1976. In the first the total period fertility rate (TPFR) was assumed to decline by around 40 per cent over a 15-year period from about 3.5 in 1976 to 2.1 in 1991. The 1991 replacement level was the same as the long-term assumption then being made in the official national projections and so it was assumed that differences in fertility would disappear by 1991. In the second assumption, the rate of fertility decline was halved to give a TPFR of 2.7 by 1991. The fertility of NCWP ethnic origin women born in Britain was assumed to fall half way between overseas born women and replacement level. The fertility of women born in NCWP who arrived in Britain after mid-1976 was assumed to be the same as the fertility of NCWP-born women in Britain in mid-1976, and then to fall in line with the decline for NCWP-born women.

Births were also projected for the population of women of wholly NCWP ethnic origin by removing children who were thought to have had fathers from the 'indigenous population'. An arbitrary assumption was made that the number of these children would double over the projection period from 1976 to 1991.

Two migration assumptions were made. The first assumed that net in-migration of those of NCWP ethnic origin would continue at about 40,000 per year over the period of the projection, that is, roughly at the average level of the five years from mid-1971 to mid-1976. The second assumed a decline to a net in-migration of 15,000 per year by 1982, remaining at that level over the rest of the projection period.

The two analyses giving lower and higher variant projections projected there would be between 2.25 and 2.5 million people of wholly or partly NCWP ethnic origin in 1986. These projections were broadly consistent with earlier assessments that the number might be of the order of 2.7 million by 1986 (OPCS, 1977), and the actual estimate derived from Labour Force Survey average of 1985–87 that gave a total minority ethnic population of 2.47 million (Haskey, 1988). (The LFS estimate includes all minority ethnic populations other than the White population, and not just the NCWP population). The author of the 1976–91 projections acknowledges that the proximity of the projections to the earlier ones is 'to some extent fortuitous' as the earlier ones were made with such limited data. In 1986 these 1976–91 projections were assessed (OPCS, 1986) and it was concluded that the 1976-based higher variant projections 'continue to give a useful guide to the prospective growth of the non-White population over the next few years'.

The 1976–91 projections gave a total 1991 population of wholly or partly NCWP ethnic origin of between 2.75 to 3.25 million. The 1991 Census recorded 2.14 million people in households headed by a non-White person born in NCWP, as well as 442,000 in households where the head was of Black or Asian ethnic origin and born in the UK (those in the 'Other group' were excluded as the majority were associated with non-New Commonwealth countries) (OPCS and General Register Office for Scotland, 1993). Added together, a very rough estimate of 2.58 million population of NCWP ethnic origin in Britain is obtained – somewhat lower than the 1976–91 projections. Using 1991 Census data and definitions of the different ethnic groups, the overall minority ethnic population was 3.015 million – which was in the middle of the range given by the 1976–91 projections. However, the 1991 Census included all the minority ethnic populations born outside the New Commonwealth, and also others born in the UK to parents who had been born in non-Commonwealth countries. The projections did not include these groups so the figures are not strictly comparable.

The results of the crude projections from 1991 to 2001 showed a projected NCWP population in Britain in 2001 of around 3–4 million, between 5 and 7 per cent of the projected overall population. These projections also included those who were *partly* of NCWP ethnic origin. The difficulties of including them were acknowledged

and the 'arbitrary' assumptions made were described. These assumptions entailed estimating not only the rates of inter-marriage between ethnic groups, but also the numbers of births and deaths and how they would change over time. The range of the results is quite large. Labour Force Survey data show that in mid-2000 the total minority ethnic private household population of Britain (not just NCWP-born) was 4.0 million or 7 per cent of the total (Scott et al, 2001).

Simple projections were published by OPCS in 1986 that examined the growth of the NCWP population between 1951 and 1984 and estimated that the growth had been around 90,000 per year (OPCS, 1986b). This rate of increase was assumed to continue and projected figures were published for 1986 (2.5 million) and 1991 (3.0 million). Straightforward estimates of the numbers of births, deaths and migrants were made, and the resulting projection was compared with that from the 1976-based projections. It was concluded from this 'tentative evidence' that the growth suggested by the 1976-based projections may have been slightly over-estimated.

The projections and estimates outlined above show that some official national projections of the NCWP population have been made, although *subnational* projections have not been attempted. Data problems were overcome with a series of reasoned assumptions made by experienced demographers. Some of the earlier estimates were examined in retrospect and the results used to refine techniques. Improved estimates were then produced, acknowledged and published. Many of the problems associated with these earlier projections would not be present today as the 1991 and 2001 Censuses provide comparatively very good estimates of the populations in each ethnic group. However there are still difficulties with fertility and mortality data, which are still recorded by birthplace and not by ethnic group, and with ethnic migration data, which are weak between census years.

1.4 Other British projections involving ethnicity

Apart from OPCS, some other organisations have undertaken population projections involving ethnicity. Bradford Council (Bradford Council, 2000) and the London Research Centre (LRC, 1999) have produced projections for many years now. Bradford Council uses a standard single year of age cohort survival model by gender from 1991 to 2021, with separate inputs for fertility, mortality, overseas migration and UK migration. It projects the populations of six ethnic groups. The assumptions for the components of fertility, mortality and migration are that past rates will continue – apart from a slow reduction in South Asian fertility, and a steady reduction in mortality as is assumed in national projections. Actual recorded numbers of births, deaths and migration are used up to 1995, which acts as the base year. Bradford have been able to calculate much more accurate rates of fertility and mortality than anywhere else in Britain as the ethnic group of each child is collected at birth – and names from deaths data have been run against Nam Pehchan, the South Asian names database, which allocates an ethnic group on the basis of the family name of the individual. This names database works well in Bradford where the majority of the ethnic minority community are from the South Asian ethnic groups. However, it would not work as well in other more ethnically diverse areas.

The London model also uses a cohort survival model projecting populations for ten ethnic groups for each of the 32 London boroughs and the City of London. Each ethnic group population is projected by single year of age and gender. The components of fertility, mortality and migration are calculated separately, together with the ethnic composition of areas affected by boundary changes. A press launch was carefully planned when the figures were first released to provide journalists with the full context, and an endorsement of the projections by the Commission for Racial Equality accompanied the figures.

More recently, Oldham (Oldham Metropolitan Borough Council, 1997) and Rochdale (Bradford Council, 1999) local authorities have produced projections. Both these projections are based on the model devised by Bradford Council. A number of other local authorities have also investigated producing projections, particularly based on a model – POPGROUP – developed in 1999 by Bradford local authority. Birmingham Council are planning to release projections by ethnic group shortly.

The strengths of these models are that they produce current and future estimates by ethnic group that are unavailable from any other data source. For some ethnic groups whose populations are changing rapidly, it is invaluable to be able to use updated estimates and projections, rather than merely relying on census

data, which may be ten years old. Local authorities have produced these projections to plan for – and provide – appropriate services for the communities they serve. The projections are very widely used. The weaknesses of the models centre on the amount of manipulation of data required in order to produce demographic rates by ethnic group. The weakest component is probably migration due to the lack of migration data by ethnic group between censuses.

Other projections include those of the Catholic and Protestant populations in Northern Ireland (Compton, 1996). Compton begins by presenting a simple extrapolation of past trends, estimated rates of natural increase between 1981 and 1991, and rates of intercensal change between 1981 and 1991, and includes some detail at district level. When these past trends are projected, they imply a Catholic majority within the next 50 years. However, more detailed evidence is then presented of decreasing fertility amongst Catholics and a plateauing out of decreasing fertility amongst non-Catholics – which suggests that a Catholic majority in the Province is not inevitable. Four projection scenarios for Northern Ireland as a whole are produced with different fertility and migration assumptions. Age-specific fertility rates are calculated for each denomination. Mortality rates are taken from the 1990–1992 Northern Ireland life table – and applied to both denominations. Net migration is taken from 1981 to 1991 intercensal change and allocated to age groups using National Health Service Central Register (NHSCR) data. The projection scenarios point to Catholic and Protestant populations becoming more finely balanced in size in the future. Fertility patterns seem to suggest that Protestants may well remain the majority in the foreseeable future, but only just. Future migration patterns are more difficult to predict, but unless new large differences emerge in migration patterns between the two groups, non-Catholics will just remain the majority group. These population projections exemplify the usefulness in identifying possible future scenarios and highlighting those thought most likely to occur. They also identify the difficulties of projecting migration patterns and, in this case, show that these assumptions can influence which group becomes the majority population – a sensitive issue in this and many other areas.

Peach (1990) derived estimates for the Bangladeshi community in Britain in 1951, 1961 and 1971 that were back projections from 1981. There were particular difficulties in estimating Bangladeshis during this period as Bangladesh obtained its independence in 1971 and therefore, before that date, ethnic Bangladeshis would have been recorded as being born in Pakistan. Peach used the Labour Force Survey to estimate the numbers of Bangladeshis who had entered Britain up to the 1980s. He then expressed the number of Bangladeshis as a proportion of the number of Pakistanis from these data, and applied this proportion to the combined population of both Pakistanis and Bangladeshis from the 1971 Census to obtain a 1971 estimate for the numbers of Bangladeshis. He also used special tabulations from the ONS Longitudinal Study to estimate Bangladeshis who were in the sample in both 1981 and 1971 to derive another 1971 estimate. He used similar techniques to obtain estimates of Bangladeshis for 1961 and 1951.

Owen (1996b) and Peloe and Rees (1999) have both produced 1981 estimates by ethnic group that are, in fact, back projections from 1991 to 1981. Owen used a component method, estimating fertility and mortality rates based on country of birth data that included assumptions about the convergence of fertility and mortality over time. Peloe and Rees calculated conditional probabilities of ethnicity given country of birth, using several different data sources. These back projections will provide a useful measure of growth between 1981 and 1991 that can then be related to growth between 1991 and 2001. Indications of growth rates for 20 years will then be available for each ethnic group if problems of slightly different categories in 2001 can be overcome.

Projections of the British minority ethnic population were undertaken by Ballard and Kalra (1994). The methodology was not fully described but did include assumptions of zero net migration and of fertility levels soon dropping to replacement levels. The report concluded that Britain's minority population was 'virtually certain' to grow very substantially in the years to come, probably to somewhere in the order of double its current size. The eventual size was predicted 'not to amount to very much more than 10 per cent of Britain's population as a whole'.

The work summarised in this section shows that models have been designed by academics and local government officers to project, at local level, the populations of a number of ethnic groups, or selected ethnic groups. Lessons can be learned from those producing these projections regarding

the data and methodologies used, as well as from consultation and dissemination strategies.

1.5 International work

Given that there has been, until now, relatively little national work in this area, it is useful to examine relevant work from other countries. A survey carried out by the Netherlands Central Bureau of Statistics, the Bureau of the Census (USA Department of Commerce) and the Netherlands Interdisciplinary Demographic Institute in early 1988 (Keilman and Cruijsen, 1992) identified five 'industrialised' countries that had produced projections by race/nationality. These were: the USA; Belgium; Federal Republic of Germany; Luxembourg; and Switzerland. The Netherlands (Statistics Netherlands, 1999), Denmark, Sweden, Canada, Germany (Munz and Ulrich, 1997), the USSR (Kingkade, 1988), Malaysia (Yusof, 1986) and France (Le Bras, 1997) have also produced projections by ethnicity/birthplace. Further race/nationality/ethnicity projections have been carried out for particular groups or regions such as Maoris in New Zealand (Pool and Pole, 1987), Aboriginal and Torres Strait Islander populations (Australian Bureau of Statistics, 1998), native Indians in Canada (Nault et al, 1995), ethnic groups in Gui Zhou in China (Wang and Guo, 1997), and for the Igbo population in Nigeria (Iro, 1989), Iraqi Kurds (Daponte et al, 1997), and gypsies in Czechoslovakia (Kalibova, 1990).

The above list illustrates the variation in the types of projections undertaken, which leads to differences in the methodology being used to undertake those projections. Firstly the list includes projections by ethnicity, birthplace, and nationality. Very few of the projections listed above project the population for each ethnic group. The USA provides one of the few exceptions, and US projections will be described in more detail below. For many of the projections, the focus is on the growth of the foreign-born population compared with the native-born population. Others, such as those for Sweden and Denmark, attempt to distinguish the foreign-born population as well as native-born descendants of immigrants. By combining the two, estimates of the minority ethnic population can be obtained. Secondly, methodological issues are highlighted from the list above, as some countries project populations of all ethnic/birthplace groups in the country, whilst others, such as the projections of Maoris in New Zealand and Aborigines in Australia, concentrate on projections for one particular ethnic group.

Most of the projections that are carried out for a range of ethnic/birthplace groups are undertaken using a component method with separate assumptions of fertility, mortality and migration. Data quality varies considerably for all the countries mentioned, but even in the data-rich countries, data weaknesses have led to problems, such as the underestimation of mortality rates for some groups in the USA and Sweden.

The USA probably has the longest tradition and the largest literature on race/ethnic group projections, and also has produced the most detailed projections by race. National projections were produced by the US Census Bureau – quarterly from January 1999 to December 2100 by single year of age to age 100, sex, race (White, Black, American Indian, and Asian and Pacific Islander) and ethnic origin (Hispanic or non-Hispanic) and nativity status (foreign or native born) (Hollman et al, 2000). Population projections are also available for each State from 1995 to 2025 (Campbell, 1996). These were computed by single year of age, sex, race and Hispanic origin, are published on the US Census Bureau website in summary form (see References) and are available in more detail from the Census Bureau. In the national projections the base is taken from 1990 Census figures updated to 1999. The unusually long projection range and age detail is of note.

Births data are available by race and Hispanic origin, and fertility rates can be calculated by single year of age using births data and population estimates. This is a great improvement on British vital registration data, which do not record births by race or ethnic group – thus making the estimation of the fertility of different ethnic groups much more difficult. Three future fertility scenarios were produced for the USA from which high, middle and low series projections were made. 'Short term' fertility assumptions from 1999 to 2025 were made using birth expectation data and 'demographic theory', which is not discussed in more detail. The longer-term assumption was that total fertility rates for each race/Hispanic origin group would converge to replacement level fertility (2.1 children per woman) in 2150. (The total fertility rate in the US in 1991 was 2.0.) This assumption may not be appropriate for England and Wales as analysis in London has shown that the fertility of many

ethnic groups is already estimated to be below replacement level.

Deaths are recorded by race but not by Hispanic or non-Hispanic ethnic origin. Age-specific death rates (ASDRs) were calculated by race but assumptions had to be made to produce ASDRs by Hispanic/non-Hispanic origin. This is still an improvement on the British vital registration system, which, as with births, does not record the ethnic group of the person who has died. In the USA projections, there is discussion of the problems with the deaths data including difficulties encountered with the reporting of race and ethnicity, and missing and unspecific data. It was acknowledged that this could mean underestimates of death rates by up to 25 per cent for some ethnic groups, which would be a very large error. However, despite these drawbacks the data were used – and would continue to be used 'until we have a stronger basis for making adjustments'. Future death rates were based on the assumption that ASDRs would converge in the year 2150. As with fertility, three sets of assumptions for death rates were formulated – a high, middle and low series.

Previously, international migration had been projected forward at a constant level. In the latest set of projections, assumptions on international in-migration are made for separate categories of flows – family reunification (mostly from Mexico and Central America), general migration needing visas, refugees, 'undocumented migration' (which presently is estimated at around 225,000 migrants per year), and potential changes in source countries. It is very unusual to see a recognition and quantification of illegal migration in government projections. Again, a high, low and middle set of rates were formulated . Out-migration rates were estimated separately. Legal out-migration was assumed to remain constant at present rates.

Projections were also undertaken in 1996 for each state by age, race, Hispanic origin and gender from 1995 to 2025. Populations were based on 1990 Census data with adjustments for misreporting or unspecified reporting. Births are recorded by non-Hispanic race groups and Hispanic ethnic origin. These were matched with populations of women to give age race/ethnic specific fertility rates (ASFRs) for each State. Existing differentials in State ASFRs were assumed to remain constant throughout the projections period. Obviously this is unlikely to happen in practice. Annual changes in projected national birth rates were applied to the individual State rates calculated in 1994 to obtain future fertility rates for each State. Mortality rates were calculated from deaths by race/ethnic origin and populations by race/ethnic origin for each State. Future patterns of change were taken from the national projection assumptions. International migration rates were calculated from 1990 Census data by age, sex and race/ethnicity. The Census included a five-year migration question – where were you living five years ago? International emigration rates were based on national projection figures allocated to each State in relation to the proportion of the foreign-born population in each State from the 1990 Census. Internal migration rates were also based on 1990 Census data giving state to state migration flows by age, sex and race/ethnic origin. Much work was undertaken to smooth and correct the rates. Projected rates were calculated using a time series approach and a different set of data – matched Internal Revenue Service tax returns.

These projections have received some criticism (Hirschman, 1998). Hirschman's main criticisms include the problems of projecting so far ahead when ethnic identity and rates of inter-marriage are changing so rapidly. Also he believes that the projected international migration rates, whilst acceptable in the short term, are unpredictable in the longer term. He states that, although there is a healthy scepticism about the limits of population projections beyond 10–20 years amongst professional users, the figures have caused alarm amongst others. There is no doubt that these criticisms have some validity, and they highlight the need to clarify the difference between the mechanical extension of a projections model versus a short term forecast that might be more able to take account of social changes such as inter-marriage between people of different ethnic groups.

Other work in the United States has identified the different components of population growth by ethnic group, enabling backwards projections of ethnic group populations.

Statistics Canada has undertaken special one-off projections by ethnicity. Projections of Registered Indians were completed in 1985, 1990, 1993 and 1998 using the Indian Register where all vital events affecting that population are recorded (Nault et al, 1995). The richness of the available data enables this projection exercise to be carried out more easily than for a British model.

The Australian Bureau of Statistics has produced a set of projections of the Aboriginal and Torres Strait Islander populations for each state and territory (Australian Bureau of Statistics, 1998). These are termed experimental projections because of the 'experimental' nature of the base populations and deficiencies in the quality of indigenous births, deaths and migration data. The projections run from 1996 to 2006. Projections are aided by the fact that Australia undertakes a population and housing census every five years, rather than the more common ten-year period (as in Britain). Weaknesses in births and deaths data were acknowledged. Births were taken from birth registrations adjusted for underenumeration. Future birth rates were assessed using births and age structure data from past Censuses. Work was undertaken to estimate the number of births where the mother was non-indigenous and the father was indigenous. National life tables were produced for the total population for 1991–96, and it was assumed that future mortality levels would remain constant, and that national rates would apply at state and territory level. Inter-state migration data were extracted from the 1991 and 1996 Censuses, which recorded migration of the indigenous and non-indigenous population. Both five-year and one-year migration data are available. Net inter-state migration levels were held constant during the projection period. The projections also included two scenarios for future changes in the 'propensity to identify'. This exercise was undertaken in order to make allowances for the changing patterns of people recording themselves as indigenous. The changing patterns include the growing number of inter-ethnic marriages and the increasing likelihood of a child of mixed origin being identified as Indigenous. Another relevant feature of these projections is that populations for one small area were not published separately, but were included in the Australia totals, highlighting one solution for dealing with projections that might be small and subject to large variation.

In 1998, Statistics New Zealand released national projections of the Maori population, and projections of the Pacific Island population – both for 1996–2051. National projections of the Asian population were also completed for 1996–2016. Nine sets of projections were produced for each of the three ethnic groups listed above. Subnational projections were published for the Maori population from 1996–2021. A low, medium and high series of these subnational projections were produced (Statistics New Zealand, 1998). The middle series assumed declining fertility and mortality with some convergence of TFRs and life expectancy to that of the total population. From 1995, births and deaths were recorded by ethnic group, where previously these had been estimated from the parents' ethnic groups. As in Australia, having these data sets proves an advantage over British projections where fertility and mortality rates by ethnic group have to be estimated.

The Dutch have produced forecasts of the 'foreign' population from 1998–2015. These includes forecasts for ten migrant groups from: Turkey; Morocco; Surinam; Netherlands Antilles/Aruba; Africa; Asia; Latin America; Indonesia; 'total not western'; and a combined group – Europe, North America, Oceania and Japan (Statistics Netherlands, 1999).

Models have been used in Sweden and Denmark that project the numbers of the 'foreign' population. These are described in more detail below. In both cases the models were developed to distinguish overseas-born from native-born ethnic minorities. Because data there are based on country of birth, estimations have been made to identify the native-born minority ethnic minority population by using the parents' country of birth as a proxy. In Britain, self-reported ethnic group obviates the need for this analysis. However the benefits of undertaking projections this way is that the different generations of migrants may be distinguished – and they may have different social policy needs in general, and particular service provision requirements.

The Danish government does not make projections of the foreign population on a regular basis. However it made two projections, one in 1990 and one in 1997. In the second, (Statistics Denmark, 1997), the number of immigrants and then the number of descendants of immigrants were projected. Immigrants were defined as people whose parents were both foreign citizens or who were both born abroad. Descendants were people born in Denmark whose parents were both immigrants or descendants of immigrants. Projections of the numbers of immigrants/descendants were published for two broad areas of origin – more and less developed countries.

Sweden has produced two sets of projections by nationality (Statistics Sweden, 1994; Swedish Immigration Board, 1997). The first, in 1994, was a simple version projecting the total foreign-born population, and the total population born in

Sweden with one or both parents born outside Sweden. The second set projected populations for nine groups. These were based on three sets of countries: the Nordic countries; the rest of Europe plus the USA, Canada, Japan, Australia and New Zealand; and thirdly the rest of the world. Projections were made for each of these groups: migrants born in those countries; those born in Sweden with both parents born outside of Sweden; and those born in Sweden with one parent born in Sweden and one parent born outside Sweden. Base populations were taken from the population register. Fertility was calculated using information from the population register but because of data weaknesses a time series of period fertility was used and the assumption made that the fertility of immigrants would gradually approach the 'Swedish' level. Mortality was also calculated from the population register but data weaknesses again led to mortality rates for the whole population being used for each group. Migration in the Swedish model was calculated from the migration register. Migration flows were assumed to increase over time, but *net* migration was assumed constant in the future.

In Switzerland, projections are not produced regularly, but a set were produced by the Federal Statistical Office in 1995 (Office Fédéral de la Statistique, 1996). Ethnic issues have not yet raised the kind of public concern and academic interest that they have in the UK. Public and political debate is focused on foreign populations and immigration policies, in view of the high percentage of foreign residents in Switzerland (Office Fédéral de la Statistique, 2000). These concerns are reflected in the data, which provide detailed analyses of foreign nationals, but do not give any information on ethnic groups.

Overall, international work on these kind of projections is patchy, but over 20 countries have produced ethnic/birthplace/citizenship projections, including most Western European countries. Very few have produced these on a regular basis. The USA is the most advanced, regularly producing projections in detail for many ethnic/race groups. It is striking that the accounts of methodology are often very weak and data sources are not described in detail but perhaps this is partly due to the fact that the accounts are often written for internal use within the country, with readers being expected to be familiar with the data.

1.6 Conclusion

This survey of literature available on ethnic group projections highlights a number of issues. The amount of literature available is limited which reflects the difficulties faced in undertaking projections of this kind. However, although projections are often produced intermittently and for specific ethnic groups, over 20 countries now have attempted ethnic/birthplace or nationality projections. Projections for Great Britain for the whole NCWP population were undertaken by OPCS from the early 1970s to the mid 1980s. Five local authorities and a few academics have also produced ethnic group projections.

There is obviously a perceived need for these figures, which has driven demographers to construct projections models despite the difficulties that exist with the quality of data everywhere. In some countries there is a lack of accurate demographic data by ethnic group, whilst in others the quality of the available data is poor. Problems therefore arise with the construction of different components of a projection model using those data. Data that have been used in local ethnic group projections models are not all available nationally. Each model has to make the best use of available data for the required geography, although methodologies can easily be shared between models.

The majority of models that have been developed are basic in their construction, with most of them simply projecting current levels of fertility, mortality and migration into the future. Some, such as the USA, London and Bradford models are more complex. Technical problems, such as changing ethnic identity over time, are also identified and tackled by attempting to project complex and fluid variables. However many producers have begun by developing simple models and have then refined the methodologies.

Finally, accounts of the methodology of many of the models tend to be very weak and would not allow the reader to replicate the models. Often data sources are not described, but this is perhaps because accounts are often written for internal users within the country or area, who are assumed to be familiar with the data sets available. It is an important lesson that the methodology should be fully documented with clear outlines of the assumptions made.

References

Australian Bureau of Statistics (1998) *Experimental projections of the Aboriginal and Torres Strait Islander population*, ABS: Canberra.

Ballard R and Kalra V S (1994) *The ethnic dimensions of the 1991 Census: a preliminary report*, Manchester Census Group, University of Manchester: Manchester.

Barton P (2000) *Ethnic minority population in Luton*. Paper submitted to Director of Planning and Development, Luton Borough Council, 15 March 2000.

Bradford Council (1999) *Population forecasts for Rochdale 1999–2021: age, sex and ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford.

Bradford Council (2000) *Population forecasts for Bradford 1999–2021: age, sex and ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford.

Campbell P (1996) *Population projections for States by age, sex, race and Hispanic origin: 1995 to 2025*, report PPL-47, US Bureau of the Census, Population Division: Washington DC. (See: www.census.gov/population/www/projections/popproj.html.)

Compton P (1996) Indigenous and older minorities, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census, volume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 243–82.

Daponte B O et al (1997) Bayesian demography: projecting the Iraqi Kurdish population, 1977–1990. *Journal of the American Statistical Association*, vol 92 no 440, pp 1,256–67.

Eversley J et al (2001) *Ethnicity and the costs to local government*. Report for HEASIG/Local Government Association, London.

Hansard (1970) *The Official Report, House of Commons*, 23 March 1970, fifth series, vol 798, cols 285–6.

Harding S, Dews H and Simpson S L (1999) The potential to identify South Asians using a computerised algorithm to classify names. *Population Trends* 97, pp 46–9.

Haskey J (1988) The ethnic minority populations of Great Britain: their size and characteristics. *Population Trends* 54, pp 29–31.

Hickman M and Walter B (1997) *Discrimination and the Irish community in Britain*, Commission for Racial Equality: London.

Hirschman C (1998) *Race and ethnic population projections: a critical evaluation of their content and meaning*. Revision of a paper presented at the 13th SUNY-Albany Conference, 1994. Seattle Population Research Center, University of Washington: Seattle.

Hollman F W et al (2000) Methodology and assumptions for the population projections of the United States: 1999 to 2100, Population Division Working Paper no 38, US Bureau of the Census: Washington DC.

Iro, M I (1989) Projecting Igbo population to the year 2000 and beyond. *Genus*, vol 45 no 3–4, pp 153–64.

Kalibova K (1990) Prognosis of the Gypsy population in Czechoslovakia up to the year 2005. *Demografie*, vol 32 no 3, pp 219–24.

Keilman N and Cruijsen H (eds) (1992) *National population forecasting in industrialized countries*, NIDI/CBGS Publication no 24, Swets and Zeitlinger: Amsterdam.

Kingkade W W (1988) *USSR: estimates and projections of the population by major nationality, 1979 to 2050*, Center for International Research Staff Paper no 41, US Bureau of the Census, Center for International Research: Washington DC.

Le Bras H (1997) Projections of the foreign population: a ten year perspective. *Population*, vol 52 no. 1, pp 103–33.

London Research Centre (1999) *1999 Round of ethnic group projections*, LRC: London.

Munz R and Ulrich R (1997) *The future growth of the foreign population in Germany: demographic projections to 2030*, Demographie Aktuell no 12, Humboldt-University: Berlin.

Nault et al (1995) Population projections of registered Indians in Canada, 1991–2015. *Cahiers Québécois de Démographie*, vol 24 no 1, pp 109–27.

Oldham Metropolitan Borough Council (1997) *Population projections for Oldham*, OMBC, Chief Executive's Department, Policy Unit: Oldham.

Office Fédéral de la Statistique (Swiss Federal Statistical Office) (1996) *Les scénarios de l'évolution démographique de la Suisse*, 1995–2050, OFS: Berne.

Office Fédéral de la Statistique (2000) Personal correspondence from Anne-Christine Wanders, responsible for population projections at the Swiss Federal Statistical Office.

OPCS (Office of Population Censuses and Surveys) (1970) Appendix K: the estimated size of the population of New Commonwealth ethnic origin. *The Registrar General's Quarterly Return for England and Wales*, no 488, pp 50–1.

OPCS (1975) Country of birth and colour 1971–4. *Population Trends* 2. pp 2–8.

OPCS (1977a) *Social Trends* 8, Central Statistical Office, HMSO. London, pp 68–9.

OPCS (1977b) New Commonwealth and Pakistani population estimates. *Population Trends* 9, pp 4–7.

OPCS (1979) Population of New Commonwealth and Pakistani ethnic origin: new projections. *Population Trends* 16, pp 22–7.

OPCS (1986a) Estimating the size of the ethnic minority populations in the 1980s. *Population Trends* 44, pp 23–7.

OPCS (1986b) Ethnic minority populations in Great Britain. *Population Trends* 46, pp 18–21.

OPCS and General Register Office for Scotland (1993) *1991 Census: report for Great Britain (part 1)*, vol 3, HMSO. London.

Owen C (1993) Using the Labour Force Survey to estimate Britain's ethnic minority populations. *Population Trends* 72, pp 18–23.

Owen D (1996a) Size structure and growth of the ethnic minority populations, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census, volume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 80–123.

Owen D (1996b) *Estimating local change in the population of minority ethnic groups*, 1981 to 1991, Working Paper no 1, Centre for Research in Ethnic Relations, University of Warwick: Warwick.

Peloe A and Rees P (1999) Estimating ethnic change in London, 1981–1991, using a variety of census data. *International Journal of Population Geography*, vol 5 no 3, pp 179–94.

Pool I and Pole N (1987) *The Maori population to 2011: demographic change and its implications*, New Zealand Demographic Society Technical Paper no 1 1987, New Zealand Demographic Society: New Zealand.

Rogers A, Little J and Raymer J (1999) Disaggregating the historical demographic sources of regional foreign-born and native-born population growth in the United States: a new method with applications. *International Journal of Population Geography*, vol 5 no 6, pp 449–75.

Scott A, Pearce D and Goldblatt P (2001) The sizes and characteristics of the minority ethnic populations of Great Britain – latest estimates. *Population Trends* 105, pp 6–15.

Sly, F (1993) Estimating Britain's ethnic minority populations using the Labour Force Survey. *Employment Gazette*, vol 101 no 9, pp 429–31.

Statistics Denmark (1997) *Statistiske Efterretninger: Befolkning og Valg*, 1997 no 16, Danmarks Statistik: Copenhagen.

Statistics Netherlands (1999) Forecasts of the foreign population 1998–2015. *Maandstatistiek van de Bevolking*, vol 47 no 4, pp 20–36.

Statistics New Zealand (1998) *Maori population projections 1996–2051; Pacific Island population projections 1996–2051; Asian population projections 1996–2016; Subnational Maori population projections 1996–2021*. Statistics New Zealand: Christchurch. (See: www.stats.govt.nz.)

Statistics Sweden (1994) *Sveriges framtida befolkning: framskrivning för åren 1994–2050*, Demografiska Rapporter, 1994 no 3, Statistiska Centralbyrån: Stockholm.

Storkey M (2000) *Methodology of the ethnic group projections 1999 Round*, London Research Centre: London.

Swedish Immigration Board (1997) *Mångfald och ursprung: rapport från ett multietniskt Sverige*, Statens Invandrarverk: Norrköping.

Wang C and Guo F (1997) Forecasting minority group populations in Gui Zhou Province. *Chinese Journal of Population Science*, vol 9 no 3, pp 203–14.

Yusof B (1986) *Population projection by ethnicity for Peninsular Malaysia 1985–2015 and some of its implications*, Cairo Demographic Centre Monograph Series no 15, CDC: Cairo.

This is a blank page

Chapter 2

Consistency in considerations for population projections by ethnic group, and the possible role of microsimulation models

Michael Murphy

2.1 Introduction

A conventional and convenient model for making population projections is to apply fertility, mortality and migration rates to a base population broken down by age and sex (Shryock et al, 1976; Government Actuary, 2000). This approach is used for the official national, marital status and subnational population projections (Shaw, 1999; Shaw and Haskey, 1999; Office for National Statistics, 1999) and also for family projections in Britain (Murphy and Wang, 1999) because:

- it is an accurate and realistic mathematical model of the real world (populations change by births and deaths, different age groups have different probabilities of dying, moving, marrying and divorcing; only women within certain age groups give birth);
- rates should move in a more predictable way than, for example, basic numbers of births and deaths;
- assumptions can be formulated more easily using models (e.g. incorporating the effects of mortality improvement or reductions in nuptiality);
- we get details of the process as an automatic by-product, such as annual numbers of births, deaths and marriages by age and sex, unlike mathematical extrapolation methods;
- there is considerable scope for using specialised knowledge in formulating assumptions (e.g. cohabitation, old age mortality or fertility change).

Such models apply transition rates (or 'jump intensities') to a base population in order to produce an updated population year-by-year based on the transitions made by cohorts, such as those born in a particular year. It is therefore sometimes referred to as the stock and transition flow, demographic component, or cohort component method.

The other main approach for projecting subpopulations is to project the overall population using the method above, and then to project the proportions within particular subgroups, finally multiplying the overall population figure by the projected proportions in the various categories. This method is used for functional population projections such as for labour force, educational participation and household projections (Bogue et al, 1993; King et al 2000). However, such methods are based on mathematical extrapolation and cannot take into account the details of subpopulation dynamics or provide numerical figures of entries into, and exits from, these groups.

The case for using a version of the component method for **ethnic** projections *ceteris paribus* would appear to be strong, but there are two main issues that need to be considered. The first is that such models are likely to be extremely demanding in terms of data requirements; since availability of data is covered in other papers, it will not be discussed further here. The second is that there are particular issues that need to be taken into account if the method is to be applied to ethnic projections, which is the focus of Section 2.2 of this paper: Section 2.3 considers how microsimulation models might be used.

2.2 Ethnic projections using cohort component methods

2.2.1 The basic model

Conventional population projections are based on an initial female population by age. Deaths are estimated by applying the appropriate female-specific mortality rates, and births of daughters by applying the female-specific fertility rates to the women of reproductive age. This gives a projected female population one time period later, and the same process may be repeated to update the population, period by period.

The numbers of men are usually obtained as a by-product of this process: typically the number of male births is set as a multiple of the number of female births, using a sex ratio at birth of 105 or 106 males per 100 females. These male births, together with older living males, are then subject to the appropriate male mortality rates and the male population is also projected.

It would be possible to formulate a model in which the numbers of males were projected by applying paternity rates to males. However, this is unsatisfactory for a number of reasons. Information on paternity is less readily and accurately available than for maternity (in Britain at present about eight per cent of births do not have any paternity information, Office for National Statistics, 2000). Even if it were possible to do so, it would produce unacceptable results if the numbers of males and females were to be projected simultaneously and independently, since the proportion of births of each sex would drift from the biologically-determined value of about 49 per cent female.[1] Thus the conventional female-dominant model has two main advantages:

- it is simple, and can be expressed in a straightforward matrix multiplication form, the Leslie Matrix (Pollard, 1973);
- it provides consistent projections for men and women without any further modification or adjustment.

However, the assumptions underlying this model are:

(1) there are no entries into the population except by birth which, by definition, occur at age zero, and no exits except by death;

(2) births depend only on the number of women in the population and their fertility rates, and are independent of the number of men.

In practice, population change depends not only on births and deaths, but also on other forms of entry into, and exit from the population under study. If this is a nationally resident population, then the only other mechanisms – apart from birth and death – are in-migration and out-migration (although an operationally useful definition of migration may be difficult, the principle is straightforward). If the population is an ethnic group resident in a particular country, then there is the additional issue of 'migration' arising from transfer between ethnic groups. Clearly, if ethnicity is self-defined, then this

1 With sex-selective abortion, it would be possible to have a different ratio, but this is not an issue in this paper.

simply depends on changes in the individuals' (or their parents') responses. The implications of such 'migrations' (from a formal point of view, one can treat this in the same way as geographic migration) will be discussed in Section 2.2.2.

The second point, (ii), concerns the role of single-sex dominant models. Simple component models cannot incorporate factors that are known to be important as proximate determinants of fertility, such as marriage, since to have a child, a woman must be sexually active and therefore to have a male partner, in most cases – but not exclusively – this will be either a husband or a cohabiting partner. Thus the availability of men might be expected to affect the proportion of women in a partnership and, since women in partnerships are much more likely to give birth, the number of births as a consequence. In the case of ethnicity, it is likely that the reported ethnicity of an infant will be determined by the ethnicity of both mother and father, so it is not possible to ignore the joint distribution of parental ethnicity without further discussion.

2.2.2 Movement between ethnic groups

When the population can move backwards and forwards between different states, the component model is referred to as a multistate population projection method. The benefits of such a model are:

- easily updateable and non-negative populations are guaranteed;
- automatically generates a set of internally consistent projections without further ad hoc adjustment such as to a control total (although as discussed later, when this approach is extended to family units, some additional constraints are required);
- allows the construction of useful measures such as the expected future number of years to be spent in each state of the system according to the state they are in at any particular age.

The population of ethnic group *A*, P^{A}_{t+1}, in Britain at time *t+1* is related to the population at time *t* by the balancing equation

$$P^{A}_{t+1} = P^{A}_{t} + B^{A}_{t,t+1} - D^{A}_{t,t+1} + Ig^{A}_{t,t+1} + Isr^{A}_{t,t+1} - Og^{A}_{t,t+1} - Osr^{A}_{t,t+1}$$

where

B is births into ethnic group in (*t, t+1*)

D is deaths to ethnic group in (*t, t+1*)

Ig is (geographical) in-migration of ethnic group in (*t, t+1*)

Isr is self-reported definitional change into ethnic group in (*t, t+1*)

Og is (geographical) out-migration of ethnic group in (*t, t+1*)

Osr is self-reported definitional change out of ethnic group in (*t, t+1*)

From these values, appropriate rates may be calculated using standard methods (Van Imhoff and Keilman, 1991). Although there is no difference in principle in how one treats geographical and self-defined movements, I distinguish the two forms of movement because international geographical migration is incorporated in the standard treatment of population projections, and therefore it is more familiar. In practice, international migration assumptions are usually set as a numerical net value, and included as an *ad hoc* adjustment to the projection process. There are a number of reasons for this: migration tends to be discussed in terms of persons, whereas fertility and mortality trends are often discussed as rates (will the TFR be 1.72 next year? Will net in-migration be 100,000 people? etc.). Also, the main interest in international migration is in the net figure since it is all that matters for population growth estimation. However, for ethnic change, everyone who is now redefined to group *A* must have been in a different group, *B*, previously. Therefore we have the consistency relationship that

$$Isr^{A(B)} = Osr^{B(A)}$$

where

X(Y) denotes destination group *X* and origin group *Y*.

The need for consistency is directly analogous to the situation for internal migration projections where, unlike international migration, such consistency requirements cannot be ignored. There is a technically sound way of ensuring consistency. Given an initial base population, a set of relevant intensities and assumptions about

the future trends in these rates, it is possible to project the population in the various ethnic groups using the multistate projection model introduced above. The benefits of a transition-based model would appear to be substantial, but in the particular context of forecasting ethnic groups, there are some additional points that need to be considered.

- What are the dimensions of the 'state space'? i.e., how many groups will be projected, clearly age, sex, and a number of ethnic groups at the minimum, but there may be cases for producing subnational projections as well (this is ignored here, but it will become clear that while this requirement presents no additional formal problems, data availability is likely to be the main issue).

- The data requirements are very substantial. A model that is used for annual forecasting single-year-of-age populations for ages 0 to 100 and over requires 102 matrices of transition rates for each sex. Even with only six ethnic groups, there are 30 possible transitions at each age between the various ethnic groups, together with the fertility and mortality transitions, so over 7,000 separate transition rates are required for each year for which a projection is made.

The multistate model, when applied to regional projections, is often referred to as a multi-regional model, which is a one-sex model. It is therefore in direct descent from the standard population model of Section 2.1, and such a model could be applied straightforwardly to the ethnicity balancing equation above if numbers of movements between all pairs of ethnic groups and appropriate denominators were available. However, in addition, it will be necessary to decide how the ethnic group of infants is established: if it is concluded that it depends upon the ethnic group of *both* parents, then a model that includes partnerships is necessary. This raises a number of technical issues, because of constraints within the system such as the numbers of men and women who marry must be equal. Therefore the specification and fitting of such models is much more complicated than the case of regional projections where individuals do not interact with each other. With models that recognise such bilateral relationships, it is particularly difficult to model events which link the situations of individuals in different cells of the model (e.g. parental marital breakdown may lead to an alteration the child's reported ethnicity, but such linkages are difficult to implement).

2.2.3 Two-sex ethnicity projections

If a child's ethnicity depends on both parents' ethnicities, then it is necessary to model partnership, except if, for example, the ethnicity of the child depended only on that of the mother, which appears implausible[2]. If the rate of unpartnered women of group A who partner men of group B is r_{A,B^F}, then the number of such partnerships formed by women of group A not in a partnership will be $P_{A^F}\, r_{A,B^F}$, where P_{A^F} is the number of women of group A at risk of entering a partnership (assuming that a woman cannot move from one partnered state to another without dissolution of the first partnership and that people cannot have multiple simultaneous partnerships). This has to equal $P_{B^M}\, r_{A,B^M}$, where M denotes male and F denotes female. The rates will be estimated independently, but , for example, if the same rates were to be applied to the relevant populations the next year, the number of partnerships would not agree. Therefore some adjustment of one or both sets of rates in the next year is necessary (Shaw, 1999; Shaw and Haskey, 1999; Murphy and Wang, 1996, 1999).

Since we would usually require some symmetry in the partnership formation process, a mechanism that reflects some form of mutual adjustment is necessary – because otherwise it would be logically possible to have more women partnering than there are eligible men or vice versa, and because we usually assume that one sex doesn't dominate the partnership formation process. One example is that used in marital status projections which involves averaging of the relevant male and female rates until agreement is reached (Van Imhoff and Keilman, 1991), such adjustments are performed independently for each type of marriage.

In a full model that includes partnership, it would also be necessary to apply appropriate partnership dissolution rates for couples with the man in group A and the woman in group B, so that when a partnership breaks down, each member reverts to the eligible state in their group. Since breakdown is an attribute of the partnership, no similar consistency issues arise as in the case of partnership formation, but an issue that will be raised later is whether age of partners or duration of the partnership is the appropriate time axis.

2 *This would also lead to anomalies in that if there were asymmetries in the partnership rates between groups A and B. For example, if females from group A were more likely to partner men in group B than the reverse, then group A would come to dominate reported ethnicity. Even if were not the case, consider the situation where the children of women in Group A were allocated 100-x per cent to Group A and x per cent to another group: presumably, x would vary with the proportion of women with partners outside group A, and so knowledge of partnership trends would be necessary to project the value of x.*

If a partnership dissolves because of the death of one partner, then the surviving partner will move from the partnered to the unpartnered state, although the category will depend on the detail required in number of partnership statuses, e.g. widowed or simply not in a partnership.

2.2.4 Counting time

In the models considered so far, there is only one time axis, which is age, and age-specific fertility and mortality rates are used. If the model is extended to include partnership, then age-specific partnership formation rates are used (which gives the number of partnerships by age for each sex). However, it would appear more natural to consider events such as partnership dissolution in relationship to the length of the partnership, rather than exclusively to the age of the partners. First, because there is more variation by duration and secondly, this is a property of the partnership. However, using duration is effectively impossible with present component methods of population projections, because it would be necessary to cross-classify by both age and duration, which would mean that the number of cells in the model would become impracticable, as well as requiring considerably more care in implementation since these two 'clocks' would have to be updated together. This issue will be revisited in Section 2.3.

So far, it has been assumed that births occur within partnerships, but this is not always the case, and, in any event, the male partner may not be the biological father. Such issues cannot be satisfactorily dealt with in the framework sketched above, and once more, this issue will be reconsidered in Section 2.3.

2.2.5 Overview of methodological considerations

A multistate model could, in principle, provide the basis for population projections by ethnic group. However, it would be highly demanding of data but it would provide an exact mathematical model of the process of population dynamics disaggregated by ethnic group. Such models have been applied to complex population systems such as to the 45 categories of family units used by the Department of Social Security without finding any numerical instability problems (Murphy and Wang, 1997). However, such applications at the national level require considerable data from sources such as surveys for which disaggregation by ethnic group would appear not to be an option.

The model sketched out above is best considered as an ideal form to which any realistic candidate model should aspire, and alternative models can be judged by how far they meet the desiderata laid out above.

The multistate model presented in more detail in the Appendix, is a cell-based (or macro-level) model in which there are appropriate transition rates between cells. Such models are relatively inflexible which has the benefit that software to apply such methods is available and known to perform well (Van Imhoff and Keilman, 1991; Van Imhoff, 1994). On the other hand, such models cannot be extended, for example, to include duration of partnership or any form of non-Markovian dependence, i.e. the transition rate for a given cell does not depend on the history of those who are currently in that cell. If the model includes married as only one undifferentiated category, then the probability of breakdown will not depend on whether it was a first or subsequent marriage, or whether the couple cohabited before marriage, which are known to be associated with different propensities to separate. It also assumes that those in a particular category are subject to the same transition probabilities irrespective of how long they have been in that category or their previous state. This Markovian assumption is not wholly valid (Land and Rogers, 1992), and it may be relaxed, for example, by making the transition rates dependent on time spent in the state (Wolf, 1988), but this leads to massive increases in the number of parameters required and it is infeasible in most cases. In the next Section I consider microsimulation.

2.3 Microsimulation models

Demographic microsimulation has been used to elucidate kinship patterns in historical, contemporary and future populations (Smith, 1987; Wachter, 1997; Wolf 1994; Zhao, 1996). Microsimulation is a Monte-Carlo approach to projection in which individuals form the units of analysis. An initial population of individuals, subdivided by sex, age and ethnic group, is input into the computer package, and these individuals age month by month; some will: marry and/or cohabit with each other; give birth; divorce; remarry; and finally die. (In recent decades, cohabitation has become increasingly important,

and it is also included in the model, Murphy, 2000.) The occurrence of these events for each person at a given age, sex, ethnic group, marital status, etc, is determined by random numbers in accordance with probabilities of occurrence specified by the user's choice of governing demographic rates. The original population comprises unrelated individuals, but, over time, full sets of links between kin are generated, as the population evolves and these are available for analysis.

I will discuss this evolution in the context of the Berkeley SOCSIM demographic microsimulation model (Watcher, 1987, Hammel, Mason and Wachter, 1990). SOCSIM is a 'closed model' simulation, in which spouses are chosen from within the simulated population list rather than created with some profile of characteristics from outside it, as is the case with an 'open model'. This means that full bilateral sets of kinship relationships, crucially including the ethnic group of parents and children, are generated for all individuals. SOCSIM produces estimates of population statistics and kin counts that incorporate the random variability built in through the Monte Carlo process.

The model includes the option of defining a number of distinct groups, between which migration is possible. For ethnicity applications these distinct groups could be ethnic groups. Therefore changes in ethnic self-identification, referred to above, can be included. Since the model is closed, partners have to be found within the existing simulation population using an algorithm to ensure a realistic distribution of spousal age differences and ethnic group distributions (the model also includes prohibitions on incest). The model includes a partnership function that generates a pool of eligibles of either sex and then partners are found from the group of eligibles by a marriage function, such as having a preferred spousal age difference. It is possible to amend this function to allow for assortative marriage in that the function may be amended to give higher priority in the search process to members of the same group.

In the basic model, if a birth occurs within a partnership, married or cohabiting, both parents are assumed to be identified (although biologically this may not be the case), including their ethnic group. But if this is not so, then in the original SOCSIM model, no information is available for the father, but the program has been modified to identify such fathers from among all males of the appropriate age (some of whom may be partners of other women). However, obtaining information on such issues will confine such models to the realm of speculation at present. Because the model is closed, it is more complex than an open model in which a partner is created when required, but as he or she does not come with any demographic background, it is impossible to investigate general kinship relationships since only one partner in a couple will usually have such information. With SOCSIM, any kinship link may be analysed because every member of the simulation population has a link to each of their partners, their parents, their youngest child, and their next older sibling through both their mother and father (if the relevant individual has ever existed). By this means, any kinship relationship through blood or marriage may be traced through living and/or dead kin.[3]

In the original model, a child is born into the same group as her or his mother (since the basic form of migration is geographical), however; it is possible to amend the program to allocate the child to any group depending on both parents' characteristics.

The whole population simulated over the period is stored as a linked list in the computer and can be processed to produce the populations that would have been alive at any date within the simulation period (Murphy, 2001a), including information only about all kin born before the relevant date. It is then possible to calculate measures including the ethnic group of children in relation to that of their parents (or, indeed, any other kin member – including grandparents etc.).

Microsimulation models are much more flexible than the macro-level ones of Section 2.2. This is their main strength, but there are considerable overheads associated with their use. First, they are typically used for research rather than operational purposes, so they are less user-friendly. Secondly, because of their flexibility, little in the way of post-processing output may be available since this will often be done on a one-off basis, whereas, the output from a macro-level program is straightforward. Their flexibility may be at the cost of having to program alterations in a high-level programming language, and, since these programs are very intensive in use of computing resources, they may include highly optimised facilities so that a high level of sophistication may be required to include new

3 Note that this not possible where a woman gives birth without an identified partner (in this analysis, any birth outside of a cohabiting or marital union), since such fathers are regarded as unknown by the program. It would be necessary for ethnicity to be allocated in such cases, but it would not be possible to derive comprehensive measures of kinship relationships, such as the proportion of different ethnic groups with surviving parents, however such problems are likely to be minor.

facilities. The SOCSIM programs are among the few demographic microsimulation applications that have been thoroughly and publicly documented, and it has recently been rewritten. We also have access to a set of extensive post processing routines for tracing out kinship connections (and post processing is often the major problem in such studies).

Base populations for microsimulation such as those needed here cannot be taken 'off the shelf' as they can for the models discussed earlier. The sorts of household survey populations – or even census ones – that are available are not suitable since although information may be collected about, for example, children outside the household, it is very rare for even minimal information to be collected about such close relatives, and much less so for other kin (see Grundy, Murphy and Shelton (1999) for a recent example of a survey that did so). In practice, base populations have to be obtained by applying historical rates to an earlier population, which means that such rates must be available, even though such methods produce good results, the availability of extended runs of ethnic-specific vital rates is likely to be problematic.

Finally, although microsimulation is widely used for policy analysis, for the conventional forecaster, the fact that such models which rely on random numbers, will produce different results every time might appear to be a drawback (although the fact that such models are inherently stochastic is actually advantageous in assessing forecasting uncertainty if the appropriate amount of stochastic noise is included in the model, Murphy, 2001b). In practice, large simulation populations are required compared with policy model applications, with consequent implications such as that they require powerful mainframe computers rather than desktop machines. National-level projections will be based on larger simulated populations of one hundred thousand or more. For such large populations, the random variability is comparatively minor (although this may be an issue for smaller ethnic groups, since microsimulation models are known to have more stochastic variability than real-world systems, although methods of dealing with this issue have been developed, Murphy, 2001b). Attention will be concentrated on population means, not on variances; aggregating the results from a number of independent statistical replications will further reduce the standard errors of these means. While a macro-level (or cell-based) model would not be subject to such stochastic variability, the complexity of the links between different degrees of kin means that macro-level models cannot be used for such purposes.

Thus, although microsimulation is the simplest possible flexible model for this project, it would require considerable further resources in order to make it a central part of any strategy to produce population projections by ethnic group.

References

Bogue D J, Arriaga E E, Anderton D L et al (1993) *Readings in population research methodology, volume 5: population models, projections and estimates*, published for United Nations Population Fund by Social Development Center: Chicago, Ill.

Government Actuary's Department (2000) *National population projections, 1998-based*, ONS series PP2 no 22, TSO: London.

Grundy E, Murphy M and Shelton N (1999) Looking beyond the household: inter-generational perspectives on living kin and contacts with kin in Great Britain. *Population Trends* 97, pp 19–27.

Hammel E A, Mason C and Wachter K W (1990) *SOCSIM II, a sociodemographic microsimulation program, rev. 1.0, operating manual*, Graduate Group in Demography Working Paper no 29, University of California, Institute of International Studies, Program in Population Research: Berkeley, California.

King D, Hayden J, Jackson R, Holmans A and Anderson D (2000) Population of households in England to 2021. *Population Trends* 99, pp 13–19.

Murphy M (2000) The evolution of cohabitation in Britain, 1960-1995. *Population Studies* 54(1), pp 43–56.

Murphy M (2001a) *Family and kinship networks in the context of aging societies*. Paper prepared for the Conference on Population Ageing in the Industrialized Countries: Challenges and Responses organised by the Committee on Population Age Structures and Public Policy of the International Union for the Scientific Study of Population (IUSSP) and the Nihon University Population Research Institute (NUPRI), Tokyo, Japan, 19–21 March 2001.

Murphy M (2001b) *Bringing behaviour back into micro-simulation: feedback mechanisms in demographic models*. Paper prepared for Workshop on Agent-based Computational Demography, Max Planck Institute for Demographic Research, February 21–23, 2001.

Murphy M and Wang D (1996) A dynamic multi-state projection model for making marital status population projections in England and Wales, in Dale A (ed), *Exploiting national census and survey data: longitudinal and partnership analyses, CCSR Occasional Paper no 10*, CCSR: Manchester, pp 103–26.

Murphy M and Wang D (1997) *An evaluation of transition-based models for forecasting family units, analytical notes no 6*, Department of Social Security, Analytical Services Division: London.

Murphy M and Wang D (1999) Forecasting British families into the 21st century, in McRae S (ed), *Changing Britain: families and households in the 1990s*, Oxford University Press: Oxford, pp 100–37.

Office for National Statistics (1999) *Subnational population projections – England*, ONS series PP3 no 10, TSO: London.

Office for National Statistics (2000) *Birth statistics, 1999*, ONS series FM1 no 28, TSO: London.

Pollard J H (1973) *Mathematical models for the growth of human populations*, Cambridge University Press: Cambridge.

Shaw C and Haskey J (1999) New estimates and projections of the population cohabiting in England and Wales. *Population Trends* 95, pp 7–17.

Shaw C (1999) 1996-based population projections by legal marital status for England and Wales. *Population Trends* 95, pp 23–32.

Shryock H S, Siegel J S and Stockwell E G (1976) *The methods and materials of demography*, Academic Press: New York.

Smith J E (1987) Simulation of kin sets and kin counts, in Bongaarts J, Burch T and Wachter K W (eds), *Family demography: methods and their application*, Clarendon Press: Oxford, pp 249–66.

Van Imhoff E (1994) *LIPRO 3.0 user's guide and tutorial*, NIDI Working Papers. 1994/1a and 1994/1b, NIDI: The Hague.

Van Imhoff E and Keilman N W (1991) *LIPRO 2.0: an application of a dynamic demographic projection model to household structure in the Netherlands*, Swets & Zeitlinger: Amsterdam.

Wachter K W (1987) Microsimulation of household cycles, in Bongaarts J, Burch T, and Wachter K W (eds), *Family Demography: methods and their application*, Clarendon Press: Oxford, pp 215–27.

Wachter K W (1997) Kinship resources for the elderly. *Philosophical Transactions of the Royal Society of London*, Series B, Biological Sciences 352 (1363), pp 1,811–7.

Wolf D A (1994) The elderly and their kin: patterns of availability and access, in Martin L G and Preston S H (eds), *Demography of aging*, National Academy Press: Washington DC, pp 146–94.

Wolf D A (1988) The multistate life table with duration-dependence. *Mathematical Population Studies*, vol 1 no 3, pp 217–45.

Zhao Z W (1996) The demographic transition in Victorian England and changes in English kinship networks. *Continuity and Change*, vol 11 no 2, pp 243–72.

Appendix: A brief summary of multistate life table models

$\underline{l}$(sex, age, time) is a row vector of dimension *d* which gives the number of people of a given sex and age at a particular time who are in each of the *d* internal states of the system, in this case the number of ethnic groups (the sum across all values of the state space gives the overall population in the sex and age category).

$\underline{M}_i$(sex, age, time) is a *d x d* dimension matrix that gives the intensities for internal events.

$\underline{M}_e$(sex, age, time) is a column vector of dimension *d* which gives the intensities for exits (in this case there is only one mode of exit, namely death, if others such as emigration were possible the vector would be replaced by an intensity matrix of suitable dimension).

$\underline{M}_b$(sex, age, time) is an *d x d* matrix of birth intensities, the (*a, b*) elements of which give the intensity that a woman in state a gives birth to a child of a given sex in state *b*.

If a new matrix $\underline{M}$(sex, age, time) is defined as

$$\underline{M}(\text{sex, age, time}) = \underline{M}_i(\text{sex, age, time}) - \text{Diag}[\underline{M}_i(\text{sex, age, time})\,\lambda^T + \underline{M}_e(\text{sex, age, time})\,\lambda^T]$$

Where λ is a row vector of dimension *d* consisting of *1s*, the formula for updating the population becomes

$$\underline{l}(\text{sex, age}+\tau, \text{time}+\tau) = \underline{l}(\text{sex, age, time})\, e^{\underline{M}(\text{sex,age,time})\tau}.$$

Generalisations of the usual life table functions may also be derived for this model, such as the vector of person years experienced by a cohort in a given state over the following *h* years is (assuming a linear change over the period)

$$\underline{L}(\text{sex, age, time}) = \tfrac{1}{2}\, h\, \{\underline{l}(\text{sex, age, time}) + \underline{l}(\text{sex, age}+h, \text{time}+h)\}$$

Further details are given in van Imhoff and Keilman (1991).

Acknowledgements

Thanks are due to the Economic and Social Research Council who partially funded this work as part of a project Evolving trends in British kin distributions and family life experience Ref. R000237076.

Chapter 3

New models for projecting UK populations by ethnic group at national and subnational scales

Philip Rees

3.1 Introduction

Haskey (2001a) has outlined the need for official projections of the United Kingdom populations by ethnic group. This paper discusses some of the key issues that need to be addressed to produce a reliable model for ethnic group projections. The paper outlines the conventional cohort-component method used for the current United Kingdom country projections, and then shows how the model can be used for separate ethnic groups. This development exposes the need to make robust estimates of fertility rates, mortality rates and international migration flows and rates by ethnicity.

However, in the long run it does not make sense to keep the projections for each ethnic group separate. The groups interact in at least two demographic ways. The first is through persons changing ethnic group identification – which follows from using a self-identification question in a census or survey to define ethnic group membership. The second way in which ethnic groups interact is through inter-marriage or informal union so that couples who form mixed ethnicity families may then have children. The paper discusses how to introduce mixed ethnicity into the demographic projections by assigning children born to parents of different ethnicities to a new mixed category. The extent to which this assigning is done can have a profound influence on the ethnic makeup of new cohorts and eventually of the whole population.

The paper sets out the features of new models for projecting ethnic group populations in a series of linked steps. First, a national projection model for the whole population is set out, based on current practice (GAD/ONS). Next, this national model is expanded to handle different ethnic groups at first by simply replicating the whole population model for each ethnic

group. Then, in Section 3.4, the model is modified by proposing to treat immigration and emigration separately rather than just through a combined net international migration assumption. In Section 3.5, it is proposed to handle migration between the different parts of the United Kingdom (England, Wales, Scotland and Northern Ireland) as inter-regional migration rather than as contributors to the net international migration terms. The projection model then becomes a multi-regional projection model, incorporating the standard features of such models, which are also embedded in the ONS England Subnational Projection (ESP) model. In Section 3.6, this model is expanded to take in a set of subnational areas in which reside the bulk of the UK's minority ethnic groups. Section 3.7 then introduces changes in ethnic self-identification into the projection model, for which new data will become available when the processing of the 2001 Census is complete. Section 3.8 introduces mixed births and mixed ethnicity into the projection model, recognising that people of different ethnicities form families and household and have children of mixed origins. The paper then reviews the difficulties and challenges such a model design would have to overcome and the degree to which consistency with existing all group national and subnational projections could be achieved through adjustment of results to official constraints. The paper concludes with a summary of the features of the proposed models and how reasonable estimates of needed inputs could be achieved.

3.2 A national projection model for the whole population

A straightforward cohort-component model is used to project the UK population in its four constituent parts: England, Wales, Scotland, and Northern Ireland (Government Actuary, 2000; Shaw, 2000; Shaw, 2001). The parts 'interact' a little in the preparation of the net migration assumptions (net flows between the four countries), but otherwise the country projections remain unconnected. The population is projected in single age detail from one mid-year to the next. A projection model has four elements: (1) a set of equations for computing survivors, (2) a set of equations for consolidating the last age, (3) a set of equations for projecting births and (4) a set of equations for computing infant survivors.

The population of age x and gender g is projected as follows:

$$K^i_{x+1g}(t+1) = s^i_{xg}(t,t+1)\,K^i_{xg}(t) + N^i_{xg}(t,t+1) \qquad (1)$$

where

$K^i_{x+1g}(t+1)$ = persons of gender g in area i aged x+1 at time t+1

$s^i_{xg}(t,t+1)$ = survival probability for persons of gender g in area i from age x at time t to age x+1 at time t+1

$K^i_{xg}(t)$ = persons of gender g in area i aged x at time t

$N^i_{xg}(t,t+1)$ = net migrants to area i of gender g and age x at time t

Survival probabilities are derived from an analysis and forecast of mortality probabilities. Net migration numbers are derived from an analysis of recent migration flows (long-term and short-term, currently heavily positive) and assumptions based on the analysis.

Separate procedures are needed for the last and first ages. Let z+ be the oldest age group reported, age z and over, so that the end of period population is

$$K^i_{z+g}(t+1) = K^i_{z+1g}(t+1) + K^i_{zg}(t+1) \qquad (2)$$

This equation achieves the necessary aggregation of the survivors in the last and last but one period-cohort to form the start population for the next interval. For example, people aged 89 become 90 at the end of an interval, while people aged 90+ survive aged 91+ at the end of the interval. Before applying the projection computations for the next interval, the two survived populations are summed to yield the start of next interval population, aged 90+.

The first age populations involve a joint process of fertility and survival. Births are projected as

$$B^i_g(t,t+1) = w_g \sum_{x=x1}^{x=x2} f^i_x(t,t+1)\; {}^1\!/_2\left[K^i_{xf}(t+1) + K^i_{xf}(t)\right] \qquad (3)$$

where

$B^i_g(t,t+1)$ = births of gender g in area i in time interval t to t+1

w_g = the proportion of births that are of gender g (usually set to recent values)

$f^i_x(t,t+1)$ = the fertility rate for area i and females aged x in time interval t to t+1.

The first fertile age is x1 (15 in the UK national projection model) and the final fertile age is x2 (49 in this case). This is a female dominant fertility model that uses period-age fertility rates. Only the female population contributes to the birth process. The births need then to be survived to age 0 and the net number of international migrants added:

$$K^i_{0g}(t+1) = s^i_{-1g}(t,t+1)\,B^i_g(t,t+1) + N^i_{-1g}(t,t+1) \quad (4)$$

where the infant period-cohort involves age transition from x=-1 (birth) to x=0 (the first age), where $s^i_{-1g}(t,t+1)$ is the probability of survival from birth to the end of the annual interval t to t+1 and $N^i_{-1g}(t,t+1)$ is the number of net in-migrants born in the annual interval t to t+1 who survive to the end.

3.3 A national projection model for ethnic groups

How can this model be implemented for separate ethnic groups? We can introduce just one extra subscript, a, indicating ethnic group into equations (1) through (4):

$$K^i_{x+1ga}(t+1) = s^i_{xga}(t,t+1)\,K^i_{xga}(t) + N^i_{xga}(t,t+1) \quad (5)$$

$$K^i_{z+ga}(t+1) = K^i_{z+1ga}(t+1) + {}_1K^i_{zga}(t+1) \quad (6)$$

$$B^i_{ga}(t,t+1) = w_g \sum_{x=x1}^{x=x2} f^i_{xa}(t,t+1) \; {}^1\!/_2\left[K^i_{xfa}(t+1) + K^i_{xfa}(t)\right] \quad (7)$$

$$K^i_{0ga}(t+1) = s^i_{-1ga}(t,t+1)\,B^i_{ga}(t,t+1) + N^i_{-1ga}(t,t+1) \quad (8)$$

To implement this model requires estimation of baseline values and, for their future forecasting: survival probabilities, net international migration, and fertility rates for each ethnic group, chosen for projection. Survival probabilities will be derived from mortality data processed through a device such as a life table.

The choice of the number and nature of ethnic groups will therefore depend in part on how well mortality, fertility and migration variables can be estimated. The ethnic groups chosen should be, as far as possible, fairly homogeneous with respect to these demographic characteristics. The broad ethnic groups for which we will have information in the 2001 Census are shown in Table 1. This grouping has the advantage of simplicity, but also some disadvantages. Knowledge of the age structures of the households containing women in equivalent groups in the 1991 Census suggests that these broad groups will be too heterogeneous. The fertility of Black African women is probably much higher than that of Black Caribbean women, and in turn higher than that of second or third generation Black British women. Similarly, within the Asian or Asian British group, Bangladeshi women have higher fertility than Pakistani women while Indian women's fertility is at replacement or below. Chinese women's fertility is probably lower than that of the Other group. These fertility differences are likely to be more important, in the short term, than any mortality differentials. A more detailed sixteen group classification is therefore suggested based on the 2001 Census ethnic question, shown in Table 2.

Table 1

Broad ethnic categories in the 2001 Census

A.	White
B.	Mixed
C.	Asian or Asian British
D.	Black or Black British
E.	Chinese or other ethnic group

How could fertility rates be estimated for these groups (or alternative classifications)? The following points can be made.

- Direct measurement using birth records classified by ethnic group would require a change in legislation to require parents to register the ethnic group of their children. Some health authorities do note ethnic group on birth occurrence records but the classifications are too general and unreliable.

Table 2

Detailed ethnic categories in the 2001 Census

A.1	White: British
A.2	White: Irish
A.3	White: Other
B.1	Mixed: White and Black Caribbean
B.2	Mixed: White and Black African
B.3	Mixed: White and Asian
B.4	Mixed: Any other Mixed background
C.1	Asian or Asian British: Indian
C.2	Asian or Asian British: Pakistani
C.3	Asian or Asian British: Bangladeshi
C.4	Asian or Asian British: Any other Asian background
D.1	Black or Black British: Caribbean
D.2	Black or Black British: African
D.3	Black or Black British: Any other Black background
E.1	Chinese or other ethnic group: Chinese
E.2	Chinese or other ethnic group: Any other

- Use of household survey or census microdata records for households and families probably provides the best source for estimates. Indirect estimation methods used in developing countries without vital registers can be used. For example, the ratio of children under five to women of child bearing age can be used with infant survival rates to estimate fertility rates. The estimates would need to be adjusted after weighting by national population estimates by ethnic group to the more reliable fertility rates for all ethnic groups combined.

- The Household Sample of Anonymised Records from the 2001 Census (and the 1991 Census for establishing trends) could provide estimates for Government Office regions and metropolitan areas. The Labour Force Survey (LFS) could be used to monitor trends, if analysis of 2001 (and 1991) SAR and LFS shows them to give comparable results.

Mortality differentials by ethnic group could not be established by direct measurement without a change in vital registration, which is unlikely. However, the ONS Longitudinal Study of England and Wales can be used to link persons recorded in the census by ethnic group, age and sex to subsequent mortality events. As mortality impacts mainly persons aged 60 and over, estimates of mortality probabilities will be affected by the extent to which members of a particular ethnic group emigrate back to their countries of origin. This emigration effect probably reduces mortality rate estimates for Black Caribbean men and women in the retirement ages (Harding, 2000).

This effect should not matter overmuch in a projection context as long as emigration is correctly estimated, as both the deceased and emigrants disappear from the national population. However, in the current national model emigration is not explicitly represented.

3.4 A national projection model with ethnic groups, immigration and emigration

The main survival and infant survival equations should therefore be modified to distinguish emigration and immigration separately. The usual approach is to use probabilities of emigration by age and sex and to introduce immigration as numbers of immigrants added to the population. So, the survival equations would become

$$K^i_{x+1ga}(t+1) = s^i_{xga}(t,t+1)K^i_{xga}(t) + e^i_{xga}(t,t+1)K^i_{xga}(t) + I^i_{x+1ga}(t,t+1) \quad (9)$$

where

$e^i_{xga}(t,t+1)$ = probability of emigration (and survival) from area i by persons of age x, gender g and ethnic group a

$I^i_{x+1ga}(t,t+1)$ = immigrants of gender g to area i in interval t to t+1, who survive in age x+1

and

$$K^i_{0ga}(t+1) = s^i_{-1ga}(t,t+1)B^i_{ga}(t) + e^i_{-1ga}(t,t+1)B^i_{ga}(t,t+1) + I^i_{-1ga}(t,t+1) \quad (10)$$

3.5 A national projection model with ethnic groups, immigration, emigration and inter-country migration

Adding immigration flows and emigration probabilities to the national projection model, however, will require a careful re-assessment of the whole population framework. The reason is

that migration can be measured in one of two ways:

- as events, which need to be handled in a *movement* model incorporating occurrence-exposure rates; or
- as shifts in location between two points in time, which need to be handled in a *transition* model incorporating transition probabilities.

Which approach should be adopted? For a national model it probably does not make much difference, but if subnational projections are to be developed, then it makes sense to use the transition approach which is compatible with the inter-area census migration data for 2000–01 and the migration data from the FHSA registers to be released post-2001 Census. The transition demographic accounting framework needs to be defined carefully before re-formulating the national projection model. Table 3 sets out demographic accounts for a four-country division of the UK, the rest of the world, and death. The account elements are specific to one period-cohort x, one gender g and one ethnic group a for a baseline period (eg the year before the 2001 Census of mid-year 2000 to mid-year 2001). The account elements each need to be estimated from available data. For example, the mortality terms would need to be estimated by applying survivorship probabilities for period-cohorts, derived from life tables, to the start of interval populations. Equivalent accounts can be defined for the infant cohort, substituting the births for starting populations. The national projection model can now be re-written using the accounts framework to incorporate inter-country migration, and to separate out the migration and mortality processes.

Table 3

Transition accounts table for four UK countries, one period-cohort, gender and ethnic group

			Survival at time t+1 in destination j						
Starting state in period	**Ending state in period**		**England** j=1	**Wales** j=2	**Scotland** j=3	**N. Ireland** j=4	**Rest of World** j=r	**Death** d	**Totals** *(all states) + (internal states)
Existence at time t	**England**	**i=1**	K_{xga}^{e1s1}	K_{xga}^{e1s2}	K_{xga}^{e1s3}	K_{xga}^{e1s4}	K_{xga}^{e1sr}	K_{xga}^{e1d*}	K_{xga}^{e1**}
	Wales	**i=2**	K_{xga}^{e2s1}	K_{xga}^{e2s2}	K_{xga}^{e2s3}	K_{xga}^{e2s4}	K_{xga}^{e2sr}	K_{xga}^{e2d*}	K_{xga}^{e2**}
	Scotland	**i=3**	K_{xga}^{e3s1}	K_{xga}^{e3s2}	K_{xga}^{e3s3}	K_{xga}^{e3s4}	K_{xga}^{e3sr}	K_{xga}^{e3d*}	K_{xga}^{e3**}
	N. Ireland	**i=4**	K_{xga}^{e4s1}	K_{xga}^{e4s2}	K_{xga}^{e4s3}	K_{xga}^{e4s4}	K_{xga}^{e4sr}	K_{xga}^{e4d*}	K_{xga}^{e4**}
	Rest of World	**i=r**	K_{xga}^{ers1}	K_{xga}^{ers2}	K_{xga}^{ers3}	K_{xga}^{ers4}	$\varnothing$	$\varnothing$	K_{xga}^{ers+}
Totals		***(all states) + (internal states)**	K_{xga}^{e*s1}	K_{xga}^{e*s2}	K_{xga}^{e*s3}	K_{xga}^{e*s4}	K_{xga}^{e+sr}	K_{xga}^{e+d*}	K_{xga}^{e**}

Notation:

K_{xga}^{eisj} *= persons in period-cohort x, gender g and ethnic group a, in existence in region i at time t, who survive in region j at time t+1 (surviving internal migrants or stayers)*

K_{xga}^{eisr} *= persons in period-cohort x, gender g and ethnic group a, in existence in region i at time t, who survive in region r at time t+1 (surviving emigrants)*

K_{xga}^{eid*} *= persons in period-cohort x, gender g and ethnic group a, in existence in region i at time t, who die in the time interval (non-survivors)*

K_{xga}^{ei**} *= persons in period-cohort x, gender g and ethnic group a, in existence in region i at time t (starting populations)*

K_{xga}^{ersj} *= persons in period-cohort x, gender g and ethnic group a, in existence in region r at time t, who survive in region j at time t+1 (surviving immigrants)*

K_{xga}^{ers+} *= persons in period-cohort x, gender g and ethnic group a, in existence in region r at time t, who survive within the internal system (total surviving immigrants)*

K_{xga}^{e*sj} *= persons in period-cohort x, gender g and ethnic group a, in existence at time t, who survive in region j at time t+1 (survivors)*

Let the survival probabilities be defined as

$$s_{xga}^{eis*} = 1 - d_{xga}^{eid*} \quad (11)$$

where mortality probabilities are defined as

$$d_{xga}^{eid*} = K_{xga}^{eid*} / K_{xga}^{ei**} \quad (12)$$

and estimated, using the usual life table relationships

$$d_{xga}^{eid*} = \left[{}_1L_{xga}^{i} - {}_1L_{x+1ga}^{i} \right] / {}_1L_{xga}^{i} \quad (13).$$

The probabilities of emigration, given survival, are defined as

$$e_{xga}^{eisr} = K_{xga}^{eisr} / K_{xga}^{eis*} \quad (14)$$

while the probabilities of migration from one internal area to another, with survival, are defined as:

$$m_{xga}^{eisj} = K_{xga}^{eisj} / K_{xga}^{eis*} \quad (15)$$

The infant period-cohort equivalents of these probabilities are defined as

$$s_{-1ga}^{bis*} = 1 - d_{-1ga}^{bid*} \quad (16)$$

$$d_{-1ga}^{bid*} = \left[{}_1L_{0ga}^{i} - {}_1l_{0ga}^{i} \right] / {}_1l_{0ga}^{i} \quad (17)$$

$$e_{-1ga}^{bisr} = K_{-1ga}^{bisr} / K_{-1ga}^{bis*} \quad (18)$$

$$m_{-1xga}^{bisj} = K_{-1ga}^{bisj} / K_{-1ga}^{bis*} \quad (19).$$

End of period populations are projected as

$$\begin{aligned} K_{x+1ga}^{e*si}(t+1) = {} & m_{xga}^{eisi}(t,t+1)\, s_{xga}^{eis*}(t,t+1) \\ & K_{xga}^{e**}(t) + \sum_{j \neq i}^{j=n} m_{xga}^{ejsi}(t,t+1) \\ & s_{xga}^{eis*}(t,t+1)\, K_{xga}^{e**}(t) \\ & + K_{xga}^{ersi}(t,t+1) \end{aligned} \quad (20)$$

The first term on the RH side of equation (20) projects the number of surviving stayers; the second term is a sum of projected surviving internal migrants and the third term is a projection of surviving immigrants from abroad.

The last age has to be consolidated from the last but one and last period-cohort

$$K_{z+ga}^{e*si}(t+1) = K_{z+1ga}^{e*si}(t+1) + K_{zga}^{e*si}(t+1) \quad (21)$$

Births are projected conventionally using period-age fertility rates:

$$\begin{aligned} K_{ga}^{bi**}(t+1) = {} & w_g \sum_{x=x1}^{x=x2} f_{xa}^{i}(t,t+1) \\ & {}^{1}/_{2}\left[K_{xfa}^{e*si}(t+1) + K_{xfa}^{ei**}(t) \right] \end{aligned} \quad (22)$$

and then infants (persons born in the time interval) are survived

$$\begin{aligned} K_{0ga}^{b*si}(t+1) = {} & m_{-1ga}^{bisi}(t,t+1)\, s_{-1ga}^{bis*}(t,t+1) \\ & K_{ga}^{bi**}(t,t+1) + \sum_{j \neq i}^{j=n} m_{-1ga}^{bjsi}(t,t+1) \\ & s_{-1ga}^{bis*}(t,t+1)\, K_{ga}^{bj**}(t,t+1) \\ & + K_{-1ga}^{brsi}(t,t+1) \end{aligned} \quad (23)$$

This re-formulation of the transition version of the multi-regional projection model separates the processes of survival from migration. In equation (20) the start of period probabilities are relevant to the particular area (i), age (x), gender (g) and ethnicity (a) and must be estimated from mortality records and assumptions using life table methods. The survivors are then multiplied by probabilities of migration conditional on survival, estimated for the relevant origins (i), destinations (i for stayers, j for migrants), age (x), gender (g) and ethnicity (a). The immigration term is handled in a slightly different way: it is assumed that the number of surviving immigrants will be projected for the country as a whole and to this are applied probabilities of location in particular destination areas within the UK, again specific to age, gender and ethnicity.

3.6 A subnational model with ethnic groups, immigration, emigration and inter-area migration

The inter-country model can be extended to become a subnational model using exactly the same accounts framework (Table 3). Instead of four country units, a larger number of subnational units could be used. Such subnational ethnic projections are carried out by the Greater London Authority (formerly the London Research Centre) for the 33 London Boroughs, for example.

For which areas should subnational projections of ethnic groups be carried out?

Two suggestions can be put forward on the basis of an analysis of the geography of ethnic groups in 1981 and 1991 (Rees and Phillips, 1996).

- A full set of main metropolitan areas and region remainders (Table 4 shows their ethnic makeup) with suitable modification to take into account adoption of Government Office Region (GORs) as the official regional units and with the addition of Northern Ireland.

- A set of nearly 70 Local Government Areas (LGAs) with above threshold percentages of minority ethnic populations. These would include mainly metropolitan districts and London boroughs but also cities such as Blackburn, Leicester, Luton, Nottingham, Preston, Slough and Watford (Table 5), all with over 10 per cent minority ethnic population.

Can reasonable estimates of the required projection inputs be made? The answer must be yes, as many of the authorities with large minority ethnic populations already make ethnic group projections. The 2001 Census will provide the necessary detailed internal migration data, and the data needed to construct fertility rates and to check mortality estimates as proposed earlier.

A general purpose computer program for projecting subnational populations in the UK has been developed by Wilson (Wilson, 2001; Wilson and Rees, 2001), which could be extended and adapted to handle ethnic groups. The program copes with 434 areas (mid-1998 LGAs in the UK) in single year of age detail.

Although it may be possible to attempt subnational projections for the 16 groups suggested earlier at the local government area scale, it would be sensible to use the five group aggregation. It is certainly feasible using a combination of 2001 Census data, the 1991–2001 Longitudinal Study, vital and migration statistics for 2000–01 to estimate inputs to a subnational ethnic model. These inputs can be partially refreshed or checked using the post-census Labour Force Survey, and for Asian groups, electoral register names analysis can be used to monitor population stocks. However, it is unlikely, even with a substantial increase in sample size, that the Labour Force Survey can provide reliable ethnic group population estimates (Haskey, 2001b).

Table 4

Ethnic population composition of metropolitan and non-metropolitan regions, mid-1991 estimate

Region	Minority ethnic population		White population	
	%	% Change 1981-91	%	% Change 1981-91
Inner London	25.8	42	74.2	-6.0
Outer London	17.1	68	82.9	-7.5
West Midlands	14.9	30	85.1	-5.7
West Yorkshire	8.3	34	91.7	-1.3
Greater Manchester	6.1	48	93.9	-3.9
East Midlands	4.9	50	95.1	3.1
Outer Metropolitan Area	4.3	37	95.7	2.1
Rest of the North West	3.1	55	96.9	1.4
South Yorkshire	2.9	49	97.1	-2.2
Outer South East	2.3	30	97.8	7.3
East Anglia	2.1	29	97.9	9.6
Central Clydeside	2.0	64	98.0	-6.8
Rest of West Midlands	2.0	-21	98.0	5.6
Merseyside	1.9	24	98.1	-5.2
Tyne and Wear	1.9	47	98.1	-2.8
Wales	1.5	31	98.5	2.5
South West	1.4	25	98.6	7.5
Rest of the North	1.0	35	99.0	-0.3
Rest of Scotland	0.9	121	99.1	0.4
Rest of Yorkshire & Humberside	0.9	-3	99.1	4.1
GREAT BRITAIN	5.6	43	94.4	0.9

Source: Rees and Phillips (1996), estimated from OPCS data from 1981 and 1991 Censuses.

3.7 Introducing changes in ethnic self-identification into the projection model

So far, the models outlined involve no interaction between the ethnic groups. After birth, no one changes from one ethnic group to another. However, because ethnic group membership in the census involves self-identification, people can change their ethnic identities over time, just as they may put forward different identities in the home from the workplace. This phenomenon has been observed in the US Census where far more new people identified themselves as American Indian in the 1990 Census than could reasonably have been accounted for in the demographic evolution of the 1980 Census stock of American Indians.

Table 5

Local authorities with greater than average minority ethnic percentages, 1991

District	Minority ethnic %	District	Minority ethnic %
INNER LONDON		WEST YORKSHIRE	
City of London	6.8	Bradford	15.9
Camden	17.7	Kirklees	10.8
Hackney	34.0	Leeds	5.9
Hammersmith & Fulham	17.6	BEDFORDSHIRE	
Haringey	29.4	Luton	20.1
Islington	18.9	North Bedfordshire 1	0.3
Kensington & Chelsea	15.4	BERKSHIRE	
Lambeth	30.6	Reading	9.4
Lewisham	22.3	Slough	28.0
Newham	42.7	BUCKINGHAMSHIRE	
Southwark	24.8	Milton Keynes	5.9
Tower Hamlets	36.1	Wycombe	8.4
Wandsworth	20.3	CAMBRIDGESHIRE	
Westminster	20.9	Cambridge	6.2
OUTER LONDON		Peterborough	7.5
Barking & Dagenham	6.8	DERBYSHIRE	
Barnet	18.7	Derby	9.8
Bexley	5.9	GLOUCESTERSHIRE	
Brent	45.2	Gloucester	5.8
Croydon	17.8	HERTFORDSHIRE	
Ealing	32.7	Watford	10.0
Enfield	14.3	KENT	
Greenwich	12.9	Gravesham	8.5
Harrow	26.6	LANCASHIRE	
Hillingdon	12.4	Blackburn	15.8
Hounslow	24.7	Burnley	5.7
Kingston upon Thames	8.7	Hyndburn	6.0
Merton	16.5	Pendle	10.4
Redbridge	21.7	Preston	10.5
Sutton	6.0	LEICESTERSHIRE	
Waltham Forest	25.9	Charnwood	6.2
GREATER MANCHESTER		Leicester	28.9
Bolton	8.4	Oadby & Wigston	8.8
Manchester	12.8	NORTHAMPTONSHIRE	
Oldham	8.8	Northampton	6.0
Rochdale	8.0	Wellingborough	7.3
WEST MIDLANDS		NOTTINGHAMSHIRE	
Birmingham	21.8	Nottingham	10.9
Coventry	12.0	OXFORDSHIRE	
Sandwell	14.9	Oxford	9.2
Walsall	9.7	WALES	
Wolverhampton	18.9	Cardiff	7.3

Source: Rees and Phillips (1996), from OPCS data.

The 2001 Census will be the second to ask an ethnic group question. Using the Longitudinal Study it will be possible to compare answers in 1991 and 2001 to the ethnic group question for the same person for just over 1 per cent of the population. From these data, probabilities of changes in ethnic identity can be derived. Some adjustment will be needed to match comparable groups (eg to infer mixed ethnicity from the detailed 34 group classification in the 1991 Census). The changes in identity will have taken place over ten years. The projection model would require annual transitions - which can be approximated by dividing by ten. The changes will probably vary by age. They will also vary by group with children of mixed ethnicity re-identifying with the ethnicity of one or other parent.

How should changes in ethnic identity be introduced into the projection model?

The easiest way would be to apply transition probabilities to the end-of-period population stocks.

Let $P_{xg}(b|a)$ be the probability that person of age x and gender g changes identity from ethnic group a to b over a year. If it were not possible to estimate these probabilities for separate ages or the age variation was random, then we would estimate these probabilities as

$$P_{xg}\left(b\middle|a\right) = P_{g}\left(b\middle|a\right) \tag{24}$$

Note that

$$\sum_{b} P_{xg}\left(b\middle|a\right) = 1 \tag{25}$$

The end of interval stocks would be altered as follows.

$$K^{e*si}_{x+1gb}(t+1) = \sum_{a} P_{xg}\left(b\middle|a\right) K^{e*si}_{x+1ga}(t+1) \tag{26}$$

How much change might occur is unknown at the moment. This element in the projection model design may be felt to be undesirable, if the changes could not be interpreted easily in terms of social mores or pressures. But it will be useful to have this process represented in the model even if the decision is taken to set all $P_{xg}(a|a)$ equal to 1 and all $P_{xg}(b|a)$ to zero where a≠b.

3.8 Introducing mixed births into the projection model

A person of a given ethnic group may marry or cohabit with a person of another. If they have children, the children acquire characteristics, biological and cultural, from both parents. This mixture of characteristics and parental origins has been recognised by adoption of a 'mixed' category in the 2001 Census ethnicity question. Such a question has long been used in the LFS.

The assignment of an ethnicity to a child of parents of different ethnicities reflects social practice and political decisions. In some societies the fact of mixture is accepted without consequences. In other societies, offspring may be assigned by others to the mother's ethnicity or to the father's ethnicity – or to the ethnicity most visible in terms of skin colour or physiognomy. So in US society, for example, children of Black and White parents tend to be assigned by society as a whole to the Black group. In Brazilian society new groups of mixed origin such as mulatto (White/Black) or mestizo (White/Indian) emerged.

In a projection model this process needs to be recognised and the exact rules or probabilities for mixture input after discussion and careful consideration.

There are two steps in the mixing process. First, the ethnicity of the child's father must be assigned. Second, the child's ethnicity must be chosen, given the ethnicity of mother and father.

Table 6 shows a matrix of birth counts, classified by both father's ethnic group and mother's for five ethnic groups. Vital registration provides some indirect information on country of birth of father and mother, but by 2001 this is unlikely to be useful given the size and age of cohorts of children and grandchildren of immigrants in the 1950s, 1960s and 1970s. An alternative source will be the Household Sample of Anonymised Records (SAR) from the 2001 Census. It will be possible to link children to mothers and fathers (where present) and make an estimate of the probability that a birth to a mother of ethnicity a has a father of ethnicity b, for the four countries. For more robust estimates and probability estimates for subnational tables, commissioned tables from the full 2001 Census microdata would

Table 6

Births classified by parents' ethnicities

			Father's ethncity						
			White	Black	Asian	Other	Mixed	Unknown	Total
			1	2	3	4	5	6	*
Mother's ethncity	White	1	K_{11}^{b*si}	K_{12}^{b*si}	K_{13}^{b*si}	K_{14}^{b*si}	K_{15}^{b*si}	K_{16}^{b*si}	K_{*1}^{b*si}
	Black	2	K_{21}^{b*si}	K_{22}^{b*si}	K_{23}^{b*si}	K_{24}^{b*si}	K_{25}^{b*si}	K_{26}^{b*si}	K_{*2}^{b*si}
	Asian	3	K_{31}^{b*si}	K_{32}^{b*si}	K_{33}^{b*si}	K_{34}^{b*si}	K_{35}^{b*si}	K_{36}^{b*si}	K_{*3}^{b*si}
	Other	4	K_{41}^{b*si}	K_{42}^{b*si}	K_{43}^{b*si}	K_{44}^{b*si}	K_{45}^{b*si}	K_{46}^{b*si}	K_{*4}^{b*si}
	Mixed	5	K_{51}^{b*si}	K_{52}^{b*si}	K_{53}^{b*si}	K_{54}^{b*si}	K_{55}^{b*si}	K_{56}^{b*si}	K_{*5}^{b*si}
	Unknown	6	K_{61}^{b*si}	K_{62}^{b*si}	K_{63}^{b*si}	K_{64}^{b*si}	K_{65}^{b*si}	K_{66}^{b*si}	K_{*6}^{b*si}
	Total	*	K_{*1}^{b*si}	K_{*2}^{b*si}	K_{*3}^{b*si}	K_{*4}^{b*si}	K_{*5}^{b*si}	K_{*6}^{b*si}	K_{**}^{b*si}

Notes:

1. K_{ab}^{b*si} *= births surviving in region i at time t+1, to mothers of ethnicity a and fathers of ethnicity b.*

2. These data could be derived from the 2001 Census in commissioned tables.

3. The unknown elements would need to be redistributed across the known elements.

be needed. These probabilities would be best applied after the births by mother's ethnicity had been projected and survived:

$$K_{0gab}^{b*si}(t+1) = P_{Ixg}^{i}\left(b\middle|a\right)K_{0ga}^{b*si}(t+1) \qquad (27)$$

where $K_{0gab}^{b*si}(t+1)$ represents the number of infants born in the time interval t to t+1 who survive in zone i at time t aged 0, of gender g, with mother's ethnic group being a and father's ethnic group being b and where $P_{Ixg}^{i}(b|a)$is the probability for infants (I) that the father is a member of ethnic group b, given that the mother is a member of ethnic group a. The probability is specific to area i and would be expected to vary widely, depending on the ethnic composition of areas.

Table 7

Child's ethnicity as a product of parents' ethnicities under different rules

A. Mother's ethnicity dominant

		Father's ethnicity				
		White	Black	Asian	Other	Mixed
Mother's ethnicity	White	White	White	White	White	White
	Black	Black	Black	Black	Black	Black
	Asian	Asian	Asian	Asian	Asian	Asian
	Other	Other	Other	Other	Other	Other
	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed

B. Father's ethnicity dominant

		Father's ethnicity				
		White	Black	Asian	Other	Mixed
Mother's ethnicity	White	White	Black	Asian	Other	Mixed
	Black	White	Black	Asian	Other	Mixed
	Asian	White	Black	Asian	Other	Mixed
	Other	White	Black	Asian	Other	Mixed
	Mixed	White	Black	Asian	Other	Mixed

C. If mother's and father's ethnicity differ, child is mixed

		Father's ethnicity				
		White	Black	Asian	Other	Mixed
Mother's ethnicity	White	White	Mixed	Mixed	Mixed	Mixed
	Black	Mixed	Black	Mixed	Mixed	Mixed
	Asian	Mixed	Mixed	Asian	Mixed	Mixed
	Other	Mixed	Mixed	Mixed	Other	Mixed
	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed

Given knowledge of both parents' ethnicities, what ethnicity should the child have? Table 7 shows the answers that result from the application of some simple rules. Panel A assumes 'mum rules OK' and assigns only the mother's ethnicity. Panel B says that 'dad rules OK' and assigns only the father's ethnicity. Panel C adopts the rule that if the child has parents of two different ethnicities it must be placed in the mixed group. In practice, none of these rules is satisfactory. It might be better then to let parents themselves decide the child's ethnicity. This is what will happen on the 2001 Census form and what has happened for some time in the LFS and also in the 1991 Census. Table 8 reorganises the cells representing mother's and father's ethnicities as rows and represents the child's ethnicity in the columns. The body of the table contains the conditional probability of a child's ethnicity given mother's and father's ethnicity, which sums to one across each row. Table 9 shows the representation of the maximum mixing rule as probabilities in the Table 8 format.

Taking the result in equation (27) we apply these probabilities to births and sum over parent's ethnicities

$$K_{0gc}^{b*si}(t+1) = \sum_{a,b} P_{0g}^{i}\left(c\middle|a,b\right)K_{0gab}^{b*si}(t+1) \qquad (28)$$

The final step is to apply the change of ethnic identity process

$$K_{0gd}^{b*si}(t+1) = \sum_{c} P_{0g}^{i}\left(d\middle|c\right)K_{0gc}^{b*si}(t+1) \qquad (29)$$

while noting that we probably will not be able to estimate the probabilities of new ethnicity d given ethnicity c for infants or children under ten years of age from the Longitudinal Study and will need to set the probability equal to 1 if d equals c, and otherwise to 0.

3.9 Constraining projections by ethnic group to all ethnic group projections

One requirement in official projections is that subgroup projections sum to national or all group projections. For example, the England Subnational Projections (ESP) are adjusted to agree with the Government Actuary and ONS England National Projections. The adjustments are reported to be minor though no details are provided. The adjustment is achieved by simply using the following equation:

$$K^{e*si'}_{x+1g}(t+1) = K^{e*si}_{x+1g}(t+1) \left[K^{e*si}_{x+1g}(t+1) / \sum_{i \in n} K^{e*si}_{x+1g}(t+1) \right] \quad (30)$$

where n refers to the whole country and the prime indicates adjustment of the LH side variable. This adjustment is inserted at the end of each projection interval in the ESP model. The adjustment of ethnic group projections at national level can be accomplished thus:

$$K^{e*sn'}_{x+1ga}(t+1) = K^{e*sn}_{x+1ga}(t+1) \left[K^{e*sn}_{x+1ga}(t+1) / \sum_{a} K^{e*sn}_{x+1ga}(t+1) \right] \quad (31)$$

However, these adjustments of ethnic group are likely to be more problematic, because the groups grow at very different rates. For example, between 1981 and 1991 the Great Britain minority ethnic population grew by an estimated 43 per cent, while the White population grew by only 1 per cent (Table 4). The sum of ethnic group projections will be greater than the all ethnic group projections, because the weights of the faster growing ethnic groups increase over time, other things being equal. After adjustment the results of the ethnic group projections will be inconsistent with the input variables (survival probabilities, fertility rates, migration probabilities, immigration flows and emigration probabilities). To accomplish the adjustments requires, for each period-cohort, full inter-area population accounts and use of iterative proportional fitting. But even this methodology will not necessarily work without further development. Assume for a particular age group that the sum of ethnic group projections exceeds the all group projection by 5 per cent. Should all groups be weighted in some way for reduction?

Table 8

The probabilities of child's ethnicity given parents' ethnicity

Mother's ethnicity, a	Father's ethnicity, b	Child's ethnicity, c White	Black	Asian	Other	Mixed	Total
		1	2	3	4	5	*
White	White	P(1|1,1)	P(2|1,1)	P(3|1,1)	P(4|1,1)	P(5|1,1)	1
White	Black	P(1|1,2)	P(2|1,2)	P(3|1,2)	P(4|1,2)	P(5|1,2)	1
White	Asian	P(1|1,3)	P(2|1,3)	P(3|1,3)	P(4|1,3)	P(5|1,3)	1
White	Other	P(1|1,4)	P(2|1,4)	P(3|1,4)	P(4|1,4)	P(5|1,4)	1
White	Mixed	P(1|1,5)	P(2|1,5)	P(3|1,5)	P(4|1,5)	P(5|1,5)	1
Black	White	P(1|2,1)	P(2|2,1)	P(3|2,1)	P(4|2,1)	P(5|2,1)	1
Black	Black	P(1|2,2)	P(2|2,2)	P(3|2,2)	P(4|2,2)	P(5|2,2)	1
Black	Asian	P(1|2,3)	P(2|2,3)	P(3|2,3)	P(4|2,3)	P(5|2,3)	1
Black	Other	P(1|2,4)	P(2|2,4)	P(3|2,4)	P(4|2,4)	P(5|2,4)	1
Black	Mixed	P(1|2,5)	P(2|2,5)	P(3|2,5)	P(4|2,5)	P(5|2,5)	1
Asian	White	P(1|3,1)	P(2|3,1)	P(3|3,1)	P(4|3,1)	P(5|3,1)	1
Asian	Black	P(1|3,2)	P(2|3,2)	P(3|3,2)	P(4|3,2)	P(5|3,2)	1
Asian	Asian	P(1|3,3)	P(2|3,3)	P(3|3,3)	P(4|3,3)	P(5|3,3)	1
Asian	Other	P(1|3,4)	P(2|3,4)	P(3|3,4)	P(4|3,4)	P(5|3,4)	1
Asian	Mixed	P(1|3,5)	P(2|3,5)	P(3|3,5)	P(4|3,5)	P(5|3,5)	1
Other	White	P(1|4,1)	P(2|4,1)	P(3|4,1)	P(4|4,1)	P(5|4,1)	1
Other	Black	P(1|4,2)	P(2|4,2)	P(3|4,2)	P(4|4,2)	P(5|4,2)	1
Other	Asian	P(1|4,3)	P(2|4,3)	P(3|4,3)	P(4|4,3)	P(5|4,3)	1
Other	Other	P(1|4,4)	P(2|4,4)	P(3|4,4)	P(4|4,4)	P(5|4,4)	1
Other	Mixed	P(1|4,5)	P(2|4,5)	P(3|4,5)	P(4|4,5)	P(5|4,5)	1
Mixed	White	P(1|5,1)	P(2|5,1)	P(3|5,1)	P(4|5,1)	P(5|5,1)	1
Mixed	Black	P(1|5,2)	P(2|5,2)	P(3|5,2)	P(4|5,2)	P(5|5,2)	1
Mixed	Asian	P(1|5,3)	P(2|5,3)	P(3|5,3)	P(4|5,3)	P(5|5,3)	1
Mixed	Other	P(1|5,4)	P(2|5,4)	P(3|5,4)	P(4|5,4)	P(5|5,4)	1
Mixed	Mixed	P(1|5,5)	P(2|5,5)	P(3|5,5)	P(4|5,5)	P(5|5,5)	1

3.10 Discussion

Many of the model features suggested here will meet with criticism, so before summarising the key features of the model it is useful to review some of these critiques.

3.10.1 Single region versus multi-region models for projection

This debate has raged ever since Rogers adapted the Leslie matrix projection model for populations resident in sets of regions (Rogers, 1966 and 1968). In essence, the model projects both out- and in-migration by applying intensities of origin-to-destination migration to origin populations at risk. The model achieves internal consistency in the sense that the number of internal (within country) in-migrants to regions must add up to the number of internal out-migrants across the set of regions used. Independent single region projections of in-migration may lead to infeasible totals that cannot be sustained by reasonable levels of out-migration. Local governments using single region projections are often keen to boost in-migration numbers: the multi-region model provides a reality check.

Assume that local authorities run single region projections and that all assume that they gain

Table 9

Probabilities of a child's ethnicity given parent's ethnicities, adopting the maximum mixing rule

Mother's ethnicity, a	Father's ethnicity, b	Child's ethnicity, c White	Black	Asian	Other	Mixed	Total
		1	2	3	4	5	*
White	White	1	0	0	0	0	1
White	Black	0	0	0	0	1	1
White	Asian	0	0	0	0	1	1
White	Other	0	0	0	0	1	1
White	Mixed	0	0	0	0	1	1
Black	White	0	0	0	0	1	1
Black	Black	0	1	0	0	0	1
Black	Asian	0	0	0	0	1	1
Black	Other	0	0	0	0	1	1
Black	Mixed	0	0	0	0	1	1
Asian	White	0	0	0	0	1	1
Asian	Black	0	0	0	0	1	1
Asian	Asian	0	0	1	0	0	1
Asian	Other	0	0	0	0	1	1
Asian	Mixed	0	0	0	0	1	1
Other	White	0	0	0	0	1	1
Other	Black	0	0	0	0	1	1
Other	Asian	0	0	0	0	1	1
Other	Other	0	0	0	1	0	1
Other	Mixed	0	0	0	0	1	1
Mixed	White	0	0	0	0	1	1
Mixed	Black	0	0	0	0	1	1
Mixed	Asian	0	0	0	0	1	1
Mixed	Other	0	0	0	0	1	1
Mixed	Mixed	0	0	0	0	1	1

through internal migration, that is, that all areas show positive net internal migration. That situation is infeasible because there is nowhere the net in-migrants can come from. The sum of the projected net internal migration flows over a set of areas which add up to the whole country must be zero. However, single region projections can all have positive net internal migration.

The logic of the multi-regional approach has now gained official acceptance in the UK. The ONS England Subnational Projection model is multi-regional. Greater London Authority borough and ethnic projection models are multi-regional. The European Commission/EUROSTAT regional projection model has had a multi-regional design in all four rounds and a majority of subnational projection models in the European Union use multi-regional models (Rees et al, 1999).

The objection normally raised to the multi-regional model is that the number of variables that need to be input (and therefore forecast) is too large. However, most applications recognise this and estimate the model inter-regional intensities from more aggregate information. The model intensity is thus something that ensures projection consistency but is itself modelled in various ways depending on information available, the size of the state space or purpose of the projection. For example, Wilson and Rees (2001) show how a spatial interaction model of inter-regional migration developed by a team of academics for the Department of the Environment, Transport and the Regions can be incorporated into a multi-regional population projection model for 100 regions.

The key lesson to draw from this discussion is that one of the tasks in preparing a subnational projection of ethnic group populations will be careful preparation and justification of the inter-regional intensities of migration. These intensities would need to be synthetically estimated from a variety of sources including the 2001 Census, the new Patient Register Migration data for Local Authorities and the Labour Force Survey.

3.10.2 Stationarity versus time trends

Because the Patient Register migration data are now available annually, it will be possible to update migration intensities mid-decade and then examine the shifts in intensities that will have occurred between 2001 and 2006, for example. It is not necessary to investigate the time trends of each intensity: examination of trends in total out-migration by ethnicity could be combined with an assumption of constancy in the age profiles of intensity. It would also be possible to investigate time trends in in-migration flows. Again use of a spatial interaction model framework would be needed to assure consistency through use of techniques such as iterative proportional fitting, for example. The work on an upgrade of the migration model for England for the Department of Transport, Local Government and the Regions (DTLR, formerly DETR) suggests that it is vital to incorporate time trends in out-migration intensities, for example (Fotheringham and Kalogirou, 2001).

3.10.3 Feasible ethnic classifications

In the paper it is suggested that the subnational ethnic classification should be less detailed than the national, mainly on pragmatic grounds of feasibility of estimation. However, a five group classification might not satisfy particular local needs. The national classification could be used at local scale but with clear statements about the degree of reliability of estimates, with publication of results only for the larger groups.

A related point concerns the need to handle carefully changes in the ethnic classification between the 1991 Census (and subsequent LFS) and the 2001 Census (and subsequent LFS). It will be necessary to (1) harmonise prior data sets to the agreed 2001 Census based ethnic classification and (2), given harmonisation, identify the nature and degree of ethnic re-identification through a comparison of individual responses in 1991 and 2001 through the Longitudinal Study.

Care will also be needed in establishing the nature of the mixed ethnicity group in the base population. Note that the model proposed does not attempt to trace the genetic contributions of the various base population groups over time to the mixed group. This is technically feasible but not socially productive as it is an attempt to perpetuate 'apartheid' thinking about ethnicity: the mixed group should be seen as a celebration of multi-cultural Britain in demographic evolution.

What might be useful and feasible is to add to the ethnic classification a simple country of birth indicator such as UK born/not UK born. This would enable different generations of migrants into the different ethnic groups to be tracked over time. The fertility model can be adapted in a fairly straightforward way by assigning the children of overseas born parents to the native UK population. Rogers, Little and Raymer (1999) have developed such a model for the United States population.

3.10.4 The need for fathers and partnerships in the model

It could be argued that, if a female dominant fertility model is to be used, then the step of identifying a partner and choosing an ethnicity for the child based on both partner's ethnicity could be dropped, with simplification of the model and direct estimation of mother's ethnicity-child's ethnicity probabilities. I think that would be wrong. Fathers are vital in the reproduction and parenting process, and should be recognised in our demographic models: otherwise we send the wrong signals to men about their significance in the family formation process. Consequently, it might be necessary to introduce another model component for what has traditionally been called 'marital status' but which should perhaps be re-labelled 'partnership status' in recognition that legal marriage is only one of the forms of couple union leading to family formation.

3.11 Conclusions

The paper has explored how a new projection model for ethnic groups could be constructed at both national and subnational scales. In this final section, I summarise the structure of the proposed models and identify the innovations required. An illustrative application is then given, full details of which are given in the Appendix.

The paper has proposed new models for projecting ethnic group populations at national (four countries), subnational (twenty metro/non-metro regions) and local (60-70 LGAS with highest minority ethnic populations). These models have the following features.

A. Baseline period

The baseline period for which the models would be calibrated should be mid-year 2000 to mid-year 2001, because this is the annual interval most closely aligned to the census.

B. Age and time intervals, ages, gender

As in the current national projections prepared by the Government Actuary's Department, the age and time intervals should be set to one year. Populations and components should be disaggregated by single years of age from 0 to 89 with a final age group of 90+. The model should represent males and females separately.

C. Ethnic classifications and spatial units

The 2001 Census will provide three ethnic classifications: the five broad categories of Table 1, the sixteen detailed categories of Table 2 and a more detailed classification of the write-in answers (White: Other; Mixed: Any other Mixed background; Asian or Asian British: Any other Asian background; Black or Black British: Any

other Black background; Chinese and other: Any other). We would recommend use of the sixteen categories for the four country projections and the twenty region projections. For LGA projections a five category projection could be implemented.

D. Treatment of external migration

National and subnational models should treat immigration and emigration separately. Estimates can be made much more easily for these gross flow components from available information.

E. Population accounting framework and multi-regional model

The projection models must be explicitly based on the transition framework, into which census migration, both internal and external, fits most readily. Internal migration information after the census, to be provided from matching records between successive FHSA register downloads, will also fit into this framework. The projection model should be a conventional multi-regional model applied to each ethnic group.

F. Ethnic re-identification

To the conventional projection model should be added a redistribution of ethnic group populations based on re-identification changes recognised through careful analysis of 1991–2001 changes tracked via the Longitudinal Study. The changes can be applied at the end of each interval before starting the next year.

G. The ethnicity of the newborn

The fertility sub-model should introduce procedures for estimating the ethnicity of fathers as well as mothers and then of assigning the child's ethnicity based on this knowledge. Careful analysis of the 2001 Census can generate the necessary probability distributions. Use of such data, which reveal how mother and fathers actually perceive their child's ethnicity is preferable to any rule-based system.

H. Population base by ethnicity

The 2001 Census counts of persons by ethnicity, gender, age and location will be adjusted for underenumeration through the One Number Census project. These counts can be adjusted to mid-year population estimates by gender, age and location based on the 2001 Census counts.

I. Fertility rate estimates by ethnicity

Birth statistics for 2000–01 and population estimates for 2000 and 2001 can be used to produce age-specific fertility rates by location for the whole population. Analysis of child-woman ratios by ethnicity in the 2001 Census can be used to generate estimated age, ethnic and location specific fertility rates which can be adjusted to be consistent with the all group fertility rates and births counts from vital registration.

J. Mortality probability estimates by ethnicity

Analysis of ethnic group mortality using the Longitudinal Study can yield estimated mortality rates which can be adjusted to be consistent with deaths by age and sex by location for 2000–01. Appropriate life tables should be developed from the mortality rates to yield the survival probabilities needed in the projection model.

K. Internal migration and immigration estimates by ethnicity

From the 2001 Census should be extracted the array of inter-area migration by ethnicity, gender and age. It may not be appropriate to estimate the migration probabilities directly from the most detailed data because of small cell numbers. Instead the probabilities can be estimated from more robust marginal (summary) distributions (see van Imhoff et al, 1997, for techniques).

L. Emigration probability estimates

The projection models outlined in the paper require estimates of gender, age and ethnicity specific probabilities of emigration. The method usually employed is to use the parallel internal out-migration probabilities and adjust them to agree with gross estimates of emigration based on the International Passenger Survey.

M. Assumptions

Trajectories of each of the inputs to the projection models need to be specified, under assumptions about trends. Analysis of trends since 1991 will be needed to inform the process of assumption adoption. We can anticipate, however, wanting to develop scenarios that assume convergence of demographic behaviour or continued differences between the ethnicities.

This paper has put forward designs for ethnic projection models at three scales: national, subnational and local. These models attempt to capture the additional real-world complexities arising from ethnic interaction. Suggestions have been made about how the model could be made operational. The Appendix provides some back-of-the-envelope, computations that illustrate how these new features work. Are robust and reliable ethnic projections feasible? This paper argues that the answer is 'yes'.

References

Fotheringham A S and Kalogirou S (2001) *Development of a migration model: analytical and practical enhancements*. A report on the modelling procedures and results connected with the research contract variation to RADS 5/9/22 (Department of Transport, Local Government and the Regions). Department of Geography, University of Newcastle: Newcastle-upon-Tyne.

Government Actuary's Department (2000) *National population projections, 1998-based*, ONS series PP2 no 22, TSO: London.

Harding S (2000) *Ethnic minority mortality differences*. Paper presented at the ONS Ethnicity Workshop, 13 June 2000.

Haskey J (2001a) Projections of the population by ethnic group: a sufficiently interesting or a definitely necessary exercise to undertake? *Population Trends* 102, pp 34–40.

Haskey J (2001b) *Minority ethnic populations – estimates and projections*. Paper presented at the LARIA Annual Conference, University of Bath, 10 April 2001.

Rees P, Kupiszewski M, Eyre H, Wilson T and Durham H (1999) *The evaluation of regional population projections for the European Union*. Final report prepared for the European Commission Directorate General XVI (Regional Policy and Cohesion) and EUROSTAT (Statistical Office of the European Communities) for EDRF study no 97/00/74/018.

Rees P and Phillips D (1996) Geographical spread: the national picture, in Ratcliffe P (ed), *Ethnicity in the 1991 Census, volume 3: social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration*, ONS, HMSO: London, pp 23–109.

Rogers A (1966) Matrix methods of population analysis. *Journal of the American Institute of Planners*, vol 32, pp 40–4.

Rogers A (1968) *Matrix analysis of inter-regional population growth and distribution*, University of California Press: Berkeley and Los Angeles.

Rogers A, Little J and Raymer J (1999) Disaggregating the historical demographic sources of regional foreign-born and native-born population growth in the United States. *International Journal of Population Geography*, vol 5 no 6, pp 449–75.

Shaw C (2000) 1998-based national population projections for the United Kingdom and constituent countries. *Population Trends* 99, pp 4–12.

Shaw C (2001) United Kingdom population trends in the 21st century. *Population Trends*, 103, pp 37–46.

Van Imhoff E, van der Gaag N, van Wissen L and Rees P (1997) The selection of internal migration models for European regions. *International Journal of Population Geography*, vol 3, pp 137–59.

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

Wilson T and Rees P (2001) *A new subnational population projection model for the United Kingdom*. Paper presented at the European Population Conference 2001, Helsinki, Finland, 6–9 June.

Appendix: Illustrative ethnic group projections, Great Britain, 2001–11

The projections presented here are meant only to illustrate how ethnic mixing can be incorporated into projections. The inputs and assumptions are very approximate. The computations still retain net international migrant flows, which we have argued should be decomposed into immigrant and emigrant flows.

Table A.1 sets out the state-space of the projections. The models proposed in the body of the paper would be much more detailed. Table A.2 develops some crude estimates of ethnic group populations in mid-2001. Table A.3 shows an assumed age-sex structure for these estimates. The data in these two tables would be replaced by 2001 Census tabulations adjusted to 2001 mid-year population estimates.

Table A.1

State space and projection parameters

Parameters	Classes
Regions = 1	1 = Great Britain
Genders = 2	1 = Males, 2 = Females
Ages = 18	1 = 0-4, 2 = 5-9, ... , 17 = 80-84, 18 = 85+
Period-cohorts = 19	0 = birth to 0-4, 1 = 0-4 to 5-9, ... , 19 = 85+ to 90+
Ethnicities = 5	1 = White, 2 = Black, 3 = Asian, 4 = Other, 5 = Mixed
Time and age interval = 5	
Start year = 2001	
End year = 2011	
Time periods = 2	1= 2001-06, 2 = 2006-11

Table A.4 then presents the projections over two five-year periods (2001–06 and 2006–11) for five ethnic groups. The different pages contain the following information.

The first five pages of Table A.4 present the survivor and births computations for 2001–06 and 2006–11 of each of the five ethnic groups. The births computations are simple female dominant calculations which are fed into the child ethnicity computations on the final two pages of Table A.4. These computations show the two stages involved: assignment of father's ethnicity and then assignment of child's ethnicity. The illustrative computations assume a high level of inter-ethnic marriage/cohabitation and the maximum assignment of 'mixed' to the offspring of mixed ethnicity couples. The final resulting births assigned to each ethnic group are transferred back to the births entries in the main survivor tables for survival of the infant period-cohort.

These calculations for Great Britain project Whites to decrease from 53,457 thousands in 2001 (91.8 per cent of the GB population) to 53,231 thousands in 2011 (89.2 per cent). The Black population increases in size from 976 thousands in 2001 (1.7 per cent) to 1,356 thousands in 2011 (2.3 per cent). The Asian population increases from 2,266 thousands in 2001 (3.9 per cent)

Table **A.2**

Ethnic group estimates for Great Britain, 2001

Ethnic Group	Mid-1991 estimate (1000s) (1)	Estimated % change 1981-91 (2)	Mid-2001 estimate (1000s) (3)	Adjusted mid-2001 estimates (4)	Guessed % mixed (5)	Estimates mixed (6)
White	53062	0.86	53518	53687	0.5	268
Black Caribbean	522	23.60	645	647	5.0	32
Black African	219	55.02	339	340	5.0	17
Black Other	188	30.92	246	247	80.0	198
Indian	877	39.68	1225	1229	0.5	6
Pakistani	500	45.21	726	728	0.1	1
Bangladeshi	172	95.43	336	337	0.1	0
Chinese	161	44.24	232	233	5.0	12
Other Asian	202	83.18	370	371	5.0	19
Other Other	302	41.21	425	427	20.0	85
Total	56206	2.54	58063	58246	638	
Broad groups:						
White	53419					
Black	987					
Asian	2287					
Other	915					
Mixed	638					

to 2,931 thousands (4.9 per cent). The Other ethnic groups grow in population from 907 thousands in 2001 (1.6 per cent) to 1,239 thousands in 2011 (2.1 per cent). The Mixed group population starts with an estimated 634 thousand in 2001 (1.1 per cent) and grows to 913 thousand by 2001 (1.5 per cent). These results should be regarded only as showing the potential for change in size of the different groups but are consistent with the changes estimated for 1981–91 by Rees and Phillips (1996). In 1981 the White population in Great Britain numbered 52,610 thousands and made up 95.9 per cent of the total population; the minority ethnic population was estimated as 2,204 thousands (4.1 per cent). By 1991, the White population was estimated to number 53,062 thousands (94.4 per cent) and the minority ethnic population some 3,144 thousands (5.6 per cent). The minority ethnic population would, on these illustrative calculations, expand to 10.8 per cent by 2011.

Table **A.3**

Assumed age-sex composition of ethnic groups in 2001

	White	Black	Asian	Other	Mixed
Males					
0-4	3.2	5.3	5.3	5.3	5.3
5-9	3.0	5.8	5.8	5.8	5.8
10-14	2.9	5.8	5.8	5.8	5.8
15-19	3.3	5.0	5.0	5.0	5.0
20-24	3.7	4.5	4.5	4.5	4.5
25-29	3.8	3.7	3.7	3.7	3.7
30-34	3.4	4.0	4.0	4.0	4.0
35-39	3.3	4.3	4.3	4.3	4.3
40-44	3.7	2.8	2.8	2.8	2.8
45-49	3.3	2.1	2.1	2.1	2.1
50-54	2.9	2.2	2.2	2.2	2.2
55-59	2.7	1.8	1.8	1.8	1.8
60-64	2.6	1.4	1.4	1.4	1.4
65-69	2.4	0.9	0.9	0.9	0.9
70-74	1.8	0.5	0.5	0.5	0.5
75-79	1.4	0.2	0.2	0.2	0.2
80-84	0.8	0.0	0.0	0.0	0.0
85+	0.3	0.0	0.0	0.0	0.0
Total	48.6	50.3	50.3	50.3	50.3
Females					
0-4	3.0	5.2	5.2	5.2	5.2
5-9	2.9	5.7	5.7	5.7	5.7
10-14	2.9	5.6	5.6	5.6	5.6
15-19	3.1	4.8	4.8	4.8	4.8
20-24	3.7	4.8	4.8	4.8	4.8
25-29	3.8	4.2	4.2	4.2	4.2
30-34	3.4	4.2	4.2	4.2	4.2
35-39	3.3	4.0	4.0	4.0	4.0
40-44	3.7	2.9	2.9	2.9	2.9
45-49	3.3	2.1	2.1	2.1	2.1
50-54	2.9	2.0	2.0	2.0	2.0
55-59	2.7	1.5	1.5	1.5	1.5
60-64	2.8	1.2	1.2	1.2	1.2
65-69	2.9	0.7	0.7	0.7	0.7
70-74	2.3	0.5	0.5	0.5	0.5
75-79	2.0	0.2	0.2	0.2	0.2
80-84	1.5	0.1	0.1	0.1	0.1
85+	1.3	0.0	0.0	0.0	0.0
Total	51.4	49.7	49.7	49.7	49.7
Persons	100.0	100.0	100.0	100.0	100.0

Table A.4

Illustrative ethnic projections

WHITES: Survivor computations

Start Age	Start pop 2001	Surv.Prob 2001-06	Survivors 2006	Net Mig 2001-06	Final Pop 2006	End Age	Start Age	Start pop 2006	Surv.Prob 2006-11	Survivors 2011	Net Mig 2006-11	Final Pop 2011	End Age
MALES							MALES						
Birth	1519	0.9946	1510.9	0.491	1511.4	0-4	Birth	1449.2	0.9946	1441.4	0.491	1441.9	0-4
0-4	1546	0.9989	1544.3	1.056	1545.4	5-9	0-4	1511.4	0.9989	1509.8	1.056	1510.8	5-9
5-9	1596	0.9993	1594.8	1.516	1596.3	10-14	5-9	1545.4	0.9993	1544.2	1.516	1545.7	10-14
10-14	1646	0.9983	1643.1	5.704	1648.8	15-19	10-14	1596.3	0.9983	1593.6	5.704	1599.3	15-19
15-19	1634	0.9960	1627.5	13.537	1641.1	20-24	15-19	1648.8	0.9960	1642.3	13.537	1655.9	20-24
20-24	1637	0.9955	1629.7	10.914	1640.6	25-29	20-24	1641.1	0.9955	1633.7	10.914	1644.7	25-29
25-29	1801	0.9949	1791.9	4.226	1796.1	30-34	25-29	1640.6	0.9949	1632.3	4.226	1636.5	30-34
30-34	2098	0.9946	2086.7	4.563	2091.2	35-39	30-34	1796.1	0.9946	1786.4	4.563	1791.0	35-39
35-39	2177	0.9921	2159.9	2.844	2162.7	40-44	35-39	2091.2	0.9921	2074.8	2.844	2077.6	40-44
40-44	1978	0.9877	1953.7	1.091	1954.8	45-49	40-44	2162.7	0.9877	2136.1	1.091	2137.2	45-49
45-49	1746	0.9823	1715.1	0.382	1715.4	50-54	45-49	1954.8	0.9823	1920.1	0.382	1920.5	50-54
50-54	1857	0.9693	1800.1	-0.169	1799.9	55-59	50-54	1715.4	0.9693	1662.9	-0.169	1662.7	55-59
55-59	1538	0.9482	1458.4	0.292	1458.7	60-64	55-59	1799.9	0.9482	1706.8	0.292	1707.0	60-64
60-64	1301	0.9175	1193.7	0.671	1194.3	65-69	60-64	1458.7	0.9175	1338.4	0.671	1339.0	65-69
65-69	1167	0.8627	1006.8	0.345	1007.1	70-74	65-69	1194.3	0.8627	1030.4	0.345	1030.7	70-74
70-74	1008	0.7767	782.9	0.100	783.0	75-79	70-74	1007.1	0.7767	782.2	0.100	782.3	75-79
75-79	787	0.6624	521.3	0.111	521.4	80-84	75-79	783.0	0.6624	518.7	0.111	518.8	80-84
80-84	469	0.5164	242.2	0.076	242.3	85-89	80-84	521.4	0.5164	269.3	0.076	269.3	85-89
85+	313	0.3148	98.5	0.002	98.5	90+	85+	340.8	0.3148	107.3	0.002	107.3	90+
Totals,0+	26299	26361.5	47.750	26409.2	Total		Total	26409.2		26330.6	47.750	26378.4	Total
FEMALES							FEMALES						
Birth	1431	0.9956	1424.3	0.418	1424.7	0-4	Birth	1364.8	0.9956	1358.8	0.418	1359.2	0-4
0-4	1463	0.9993	1462.0	0.127	1462.1	5-9	0-4	1424.7	0.9993	1423.7	0.127	1423.8	5-9
5-9	1512	0.9996	1511.3	1.060	1512.4	10-14	5-9	1462.1	0.9996	1461.5	1.060	1462.5	10-14
10-14	1573	0.9991	1571.5	7.902	1579.4	15-19	10-14	1512.4	0.9991	1511.0	7.902	1518.9	15-19
15-19	1548	0.9985	1545.7	17.152	1562.9	20-24	15-19	1579.4	0.9985	1577.1	17.152	1594.3	20-24
20-24	1552	0.9984	1549.6	9.944	1559.5	25-29	20-24	1562.9	0.9984	1560.4	9.944	1570.4	25-29
25-29	1688	0.9978	1684.3	5.448	1689.8	30-34	25-29	1559.5	0.9978	1556.1	5.448	1561.6	30-34
30-34	1991	0.9967	1984.5	2.126	1986.6	35-39	30-34	1689.8	0.9967	1684.2	2.126	1686.3	35-39
35-39	2097	0.9948	2086.1	1.919	2088.0	40-44	35-39	1986.6	0.9948	1976.3	1.919	1978.2	40-44
40-44	1932	0.9919	1916.4	2.315	1918.7	45-49	40-44	2088.0	0.9919	2071.2	2.315	2073.5	45-49
45-49	1745	0.9873	1722.9	0.432	1723.3	50-54	45-49	1918.7	0.9873	1894.4	0.432	1894.8	50-54
50-54	1876	0.9797	1837.9	0.109	1838.0	55-59	50-54	1723.3	0.9797	1688.3	0.109	1688.4	55-59
55-59	1581	0.9678	1530.2	-0.397	1529.8	60-64	55-59	1838.0	0.9678	1778.9	-0.397	1778.5	60-64
60-64	1367	0.9496	1298.1	-0.383	1297.7	65-69	60-64	1529.8	0.9496	1452.6	-0.383	1452.2	65-69
65-69	1279	0.9125	1167.1	-0.187	1166.9	70-74	65-69	1297.7	0.9125	1184.2	-0.187	1184.0	70-74
70-74	1216	0.8526	1036.7	-0.163	1036.6	75-79	70-74	1166.9	0.8526	994.9	-0.163	994.7	75-79
75-79	1099	0.7618	837.2	-0.030	837.2	80-84	75-79	1036.6	0.7618	789.7	-0.030	789.6	80-84
80-84	806	0.6188	498.7	-0.021	498.7	85-89	80-84	837.2	0.6188	518.0	-0.021	518.0	85-89
85+	833	0.3929	327.2	-0.021	327.2	90+	85+	825.9	0.3929	324.5	-0.021	324.5	90+
Total	27158		26991.8	47.750	27039.5	Total	Total	27039.5	26805.7	47.750	26853.4	Total	

WHITES: Births to Mothers

Age at maternity	Female start pop 2001	Female end pop 2006	Pop at risk 2001-06	fert rates	births	Age at maternity	Female start pop 2006	Female end pop 2011	pop at risk	fert rates	births
15-19	1548	1579.4	1564	142	222	15-19	1579.4	1518.9	1549	142	219
20-24	1552	1562.9	1557	361	562	20-24	1562.9	1594.3	1579	361	570
25-29	1688	1559.5	1624	518	841	25-29	1559.5	1570.4	1565	518	811
30-34	1991	1689.8	1840	446	821	30-34	1689.8	1561.6	1626	446	726
35-39	2097	1986.6	2042	195	398	35-39	1986.6	1686.3	1836	195	358
40-44	1932	2088.0	2010	36	72	40-44	2088.0	1978.2	2033	36	73
45-49	1745	1918.7	1832	2	4	45-49	1918.7	2073.5	1996	2	4
Total				1700	2920	total				1700	2761

Table A.4 cont.

Illustrative ethnic projections

BLACKS: Survivor computations

Start Age	Start pop 2001	Surv.Prob 2001-06	Survivors 2006	Net Mig 2001-06	Final Pop 2006	End Age	Start Age	Start pop 2006	Surv.Prob 2006-11	Survivors 2011	Net Mig 2006-11	Final Pop 2011	End Age
MALES							MALES						
Birth	55	0.9946	54.7	0.491	55.2	0-4	Birth	59.1	0.9946	58.8	0.491	59.3	0-4
0-4	47	0.9989	46.9	1.056	48.0	5-9	0-4	55.2	0.9989	55.1	1.056	56.2	5-9
5-9	57	0.9993	57.0	1.516	58.5	10-14	5-9	48.0	0.9993	48.0	1.516	49.5	10-14
10-14	61	0.9983	60.9	5.704	66.6	15-19	10-14	58.5	0.9983	58.4	5.704	64.1	15-19
15-19	45	0.9960	44.8	13.537	58.4	20-24	15-19	66.6	0.9960	66.3	13.537	79.9	20-24
20-24	36	0.9955	35.8	10.914	46.8	25-29	20-24	58.4	0.9955	58.1	10.914	69.0	25-29
25-29	33	0.9949	32.8	4.226	37.1	30-34	25-29	46.8	0.9949	46.5	4.226	50.7	30-34
30-34	45	0.9946	44.8	4.563	49.3	35-39	30-34	37.1	0.9946	36.9	4.563	41.4	35-39
35-39	52	0.9921	51.6	2.844	54.4	40-44	35-39	49.3	0.9921	48.9	2.844	51.8	40-44
40-44	28	0.9877	27.7	1.091	28.7	45-49	40-44	54.4	0.9877	53.8	1.091	54.9	45-49
45-49	21	0.9823	20.6	0.382	21.0	50-54	45-49	28.7	0.9823	28.2	0.382	28.6	50-54
50-54	26	0.9693	25.2	-0.169	25.0	55-59	50-54	21.0	0.9693	20.4	-0.169	20.2	55-59
55-59	19	0.9482	18.0	0.292	18.3	60-64	55-59	25.0	0.9482	23.7	0.292	24.0	60-64
60-64	13	0.9175	11.9	0.671	12.6	65-69	60-64	18.3	0.9175	16.8	0.671	17.5	65-69
65-69	8	0.8627	6.9	0.345	7.2	70-74	65-69	12.6	0.8627	10.9	0.345	11.2	70-74
70-74	5	0.7767	3.9	0.100	4.0	75-79	70-74	7.2	0.7767	5.6	0.100	5.7	75-79
75-79	2	0.6624	1.3	0.111	1.4	80-84	75-79	4.0	0.6624	2.6	0.111	2.7	80-84
80-84	0	0.5164	0.0	0.076	0.1	85-89	80-84	1.4	0.5164	0.7	0.076	0.8	85-89
85+	0	0.3148	0.0	0.002	0.0	90+	85+	0.1	0.3148	0.0	0.002	0.0	90+
Total	498		544.9	47.750	592.6	Total	Total	592.6		639.8	47.750	687.6	Total
FEMALES							FEMALES						
Birth	52	0.9956	51.5	0.418	52.0	0-4	Birth	55.7	0.9956	55.4	0.418	55.9	0-4
0-4	47	0.9993	47.0	0.127	47.1	5-9	0-4	52.0	0.9993	51.9	0.127	52.0	5-9
5-9	55	0.9996	55.0	1.060	56.0	10-14	5-9	47.1	0.9996	47.1	1.060	48.1	10-14
10-14	56	0.9991	55.9	7.902	63.8	15-19	10-14	56.0	0.9991	56.0	7.902	63.9	15-19
15-19	44	0.9985	43.9	17.152	61.1	20-24	15-19	63.8	0.9985	63.8	17.152	80.9	20-24
20-24	37	0.9984	36.9	9.944	46.9	25-29	20-24	61.1	0.9984	61.0	9.944	70.9	25-29
25-29	34	0.9978	33.9	5.448	39.4	30-34	25-29	46.9	0.9978	46.8	5.448	52.2	30-34
30-34	45	0.9967	44.9	2.126	47.0	35-39	30-34	39.4	0.9967	39.2	2.126	41.4	35-39
35-39	46	0.9948	45.8	1.919	47.7	40-44	35-39	47.0	0.9948	46.7	1.919	48.7	40-44
40-44	28	0.9919	27.8	2.315	30.1	45-49	40-44	47.7	0.9919	47.3	2.315	49.6	45-49
45-49	21	0.9873	20.7	0.432	21.2	50-54	45-49	30.1	0.9873	29.7	0.432	30.1	50-54
50-54	24	0.9797	23.5	0.109	23.6	55-59	50-54	21.2	0.9797	20.7	0.109	20.8	55-59
55-59	16	0.9678	15.5	-0.397	15.1	60-64	55-59	23.6	0.9678	22.9	-0.397	22.5	60-64
60-64	11	0.9496	10.4	-0.383	10.1	65-69	60-64	15.1	0.9496	14.3	-0.383	13.9	65-69
65-69	6	0.9125	5.5	-0.187	5.3	70-74	65-69	10.1	0.9125	9.2	-0.187	9.0	70-74
70-74	5	0.8526	4.3	-0.163	4.1	75-79	70-74	5.3	0.8526	4.5	-0.163	4.3	75-79
75-79	2	0.7618	1.5	-0.030	1.5	80-84	75-79	4.1	0.7618	3.1	-0.030	3.1	80-84
80-84	1	0.6188	0.6	-0.021	0.6	85-89	80-84	1.5	0.6188	0.9	-0.021	0.9	85-89
85+	0	0.3929	0.0	-0.021	0.0	90+	85+	0.6	0.3929	0.2	-0.021	0.2	90+
Total	478		524.7	47.750	572.4	Total	Total	572.4		620.8	47.750	668.6	Total

BLACKS: Births to mothers

Age at maternity	Female start pop 2001	Female end pop 2006	Pop at risk 2001-06	fert rates	births	Age at maternity	Female start pop 2006	Female end pop 2011	pop at risk	fert rates	births
15-19	44	63.8	54	200	11	15-19	63.8	63.9	64	200	13
20-24	37	61.1	49	510	25	20-24	61.1	80.9	71	510	36
25-29	34	46.9	40	732	30	25-29	46.9	70.9	59	732	43
30-34	45	39.4	42	630	27	30-34	39.4	52.2	46	630	29
35-39	46	47.0	46	275	13	35-39	47.0	41.4	44	275	12
40-44	28	47.7	38	51	2	40-44	47.7	48.7	48	51	2
45-49	21	30.1	26	3	0	45-49	30.1	49.6	40	3	0
total				2400	107	total				2400	136

Table A.4 cont.

Illustrative ethnic projections

ASIANS: Survivor computations

Start Age	Start pop 2001	Surv.Prob 2001-06	Survivors 2006	Net Mig 2001-06	Final Pop 2006	End Age
MALES						
Birth	141	0.9946	140.4	0.491	140.9	0-4
0-4	109	0.9989	108.9	1.056	109.9	5-9
5-9	132	0.9993	131.9	1.516	133.4	10-14
10-14	141	0.9983	140.8	5.704	146.5	15-19
15-19	106	0.9960	105.6	13.537	119.1	20-24
20-24	85	0.9955	84.6	10.914	95.5	25-29
25-29	75	0.9949	74.6	4.226	78.8	30-34
30-34	105	0.9946	104.4	4.563	109.0	35-39
35-39	121	0.9921	120.0	2.844	122.9	40-44
40-44	64	0.9877	63.2	1.091	64.3	45-49
45-49	48	0.9823	47.1	0.382	47.5	50-54
50-54	60	0.9693	58.2	-0.169	58.0	55-59
55-59	44	0.9482	41.7	0.292	42.0	60-64
60-64	30	0.9175	27.5	0.671	28.2	65-69
65-69	19	0.8627	16.4	0.345	16.7	70-74
70-74	12	0.7767	9.3	0.100	9.4	75-79
75-79	5	0.6624	3.3	0.111	3.4	80-84
80-84	0	0.5164	0.0	0.076	0.1	85-89
85+	0	0.3148	0.0	0.002	0.0	90+
Total	1156		1278.1	47.750	1325.8	Total
FEMALES						
Birth	133	0.9956	132.4	0.418	132.8	0-4
0-4	109	0.9993	108.9	0.127	109.1	5-9
5-9	127	0.9996	126.9	1.060	128.0	10-14
10-14	130	0.9991	129.9	7.902	137.8	15-19
15-19	103	0.9985	102.8	17.152	120.0	20-24
20-24	86	0.9984	85.9	9.944	95.8	25-29
25-29	80	0.9978	79.8	5.448	85.3	30-34
30-34	105	0.9967	104.7	2.126	106.8	35-39
35-39	108	0.9948	107.4	1.919	109.4	40-44
40-44	65	0.9919	64.5	2.315	66.8	45-49
45-49	48	0.9873	47.4	0.432	47.8	50-54
50-54	56	0.9797	54.9	0.109	55.0	55-59
55-59	37	0.9678	35.8	-0.397	35.4	60-64
60-64	25	0.9496	23.7	-0.383	23.4	65-69
65-69	13	0.9125	11.9	-0.187	11.7	70-74
70-74	11	0.8526	9.4	-0.163	9.2	75-79
75-79	5	0.7618	3.8	-0.030	3.8	80-84
80-84	2	0.6188	1.2	-0.021	1.2	85-89
85+	0	0.3929	0.0	-0.021	0.0	90+
Total	1110		1231.3	47.750	1279.1	Total

ASIANS: Births to mothers

Age at maternity	Female start pop 2001	Female end pop 2006	Pop at risk 2001-06	fert rates	births
15-19	103	137.8	120	233	28
20-24	86	120.0	103	595	61
25-29	80	95.8	88	853	75
30-34	105	85.3	95	735	70
35-39	108	106.8	107	321	34
40-44	65	109.4	87	59	5
45-49	48	66.8	57	3	0
total				2800	274

Start Age	Start pop 2006	Surv.Prob 2006-11	Survivors 2011	Net Mig 2006-11	Final Pop 2011	End Age
MALES						
Birth	141.2	0.9946	140.4	0.491	140.9	0-4
0-4	140.9	0.9989	140.8	1.056	141.8	5-9
5-9	109.9	0.9993	109.9	1.516	111.4	10-14
10-14	133.4	0.9983	133.2	5.704	138.9	15-19
15-19	146.5	0.9960	145.9	13.537	159.4	20-24
20-24	119.1	0.9955	118.6	10.914	129.5	25-29
25-29	95.5	0.9949	95.1	4.226	99.3	30-34
30-34	78.8	0.9946	78.4	4.563	83.0	35-39
35-39	109.0	0.9921	108.1	2.844	111.0	40-44
40-44	122.9	0.9877	121.4	1.091	122.5	45-49
45-49	64.3	0.9823	63.2	0.382	63.5	50-54
50-54	47.5	0.9693	46.1	-0.169	45.9	55-59
55-59	58.0	0.9482	55.0	0.292	55.3	60-64
60-64	42.0	0.9175	38.5	0.671	39.2	65-69
65-69	28.2	0.8627	24.3	0.345	24.7	70-74
70-74	16.7	0.7767	13.0	0.100	13.1	75-79
75-79	9.4	0.6624	6.2	0.111	6.4	80-84
80-84	3.4	0.5164	1.8	0.076	1.8	85-89
85+	0.1	0.3148	0.0	0.002	0.0	90+
Total	1325.8		1439.8	47.750	1487.6	Total
FEMALES						
Birth	133.0	0.9956	132.4	0.418	132.8	0-4
0-4	132.8	0.9993	132.7	0.127	132.8	5-9
5-9	109.1	0.9996	109.0	1.060	110.1	10-14
10-14	128.0	0.9991	127.9	7.902	135.8	15-19
15-19	137.8	0.9985	137.6	17.152	154.7	20-24
20-24	120.0	0.9984	119.8	9.944	129.8	25-29
25-29	95.8	0.9978	95.6	5.448	101.0	30-34
30-34	85.3	0.9967	85.0	2.126	87.1	35-39
35-39	106.8	0.9948	106.2	1.919	108.1	40-44
40-44	109.4	0.9919	108.5	2.315	110.8	45-49
45-49	66.8	0.9873	65.9	0.432	66.4	50-54
50-54	47.8	0.9797	46.9	0.109	47.0	55-59
55-59	55.0	0.9678	53.2	-0.397	52.8	60-64
60-64	35.4	0.9496	33.6	-0.383	33.2	65-69
65-69	23.4	0.9125	21.3	-0.187	21.1	70-74
70-74	11.7	0.8526	10.0	-0.163	9.8	75-79
75-79	9.2	0.7618	7.0	-0.030	7.0	80-84
80-84	3.8	0.6188	2.3	-0.021	2.3	85-89
85+	1.2	0.3929	0.5	-0.021	0.4	90+
Total	1279.1		1395.4	47.750	1443.1	Total

Age at maternity	Female start pop 2006	Female end pop 2011	pop at risk	fert rates	births
15-19	137.8	135.8	137	233	32
20-24	120.0	154.7	137	595	82
25-29	95.8	129.8	113	853	96
30-34	85.3	101.0	93	735	68
35-39	106.8	87.1	97	321	31
40-44	109.4	108.1	109	59	6
45-49	66.8	110.8	89	3	0
total				2800	316

Table A.4 cont.

Illustrative ethnic projections

OTHERS: Survivor computations

Start Age	Start pop 2001	Surv.Prob 2001-06	Survivors 2006	Net Mig 2001-06	Final Pop 2006	End Age	Start Age	Start pop 2006	Surv.Prob 2006-11	Survivors 2011	Net Mig 2006-11	Final Pop 2011	End Age
MALES							MALES						
Birth	44	0.9946	43.8	0.491	44.3	0-4	Birth	44.1	0.9946	43.8	0.491	44.3	0-4
0-4	43	0.9989	43.0	1.056	44.0	5-9	0-4	44.3	0.9989	44.3	1.056	45.3	5-9
5-9	53	0.9993	53.0	1.516	54.5	10-14	5-9	44.0	0.9993	44.0	1.516	45.5	10-14
10-14	56	0.9983	55.9	5.704	61.6	15-19	10-14	54.5	0.9983	54.4	5.704	60.1	15-19
15-19	43	0.9960	42.8	13.537	56.4	20-24	15-19	61.6	0.9960	61.4	13.537	74.9	20-24
20-24	34	0.9955	33.8	10.914	44.8	25-29	20-24	56.4	0.9955	56.1	10.914	67.0	25-29
25-29	30	0.9949	29.8	4.226	34.1	30-34	25-29	44.8	0.9949	44.5	4.226	48.8	30-34
30-34	43	0.9946	42.8	4.563	47.3	35-39	30-34	34.1	0.9946	33.9	4.563	38.5	35-39
35-39	48	0.9921	47.6	2.844	50.5	40-44	35-39	47.3	0.9921	47.0	2.844	49.8	40-44
40-44	26	0.9877	25.7	1.091	26.8	45-49	40-44	50.5	0.9877	49.8	1.091	50.9	45-49
45-49	19	0.9823	18.7	0.382	19.0	50-54	45-49	26.8	0.9823	26.3	0.382	26.7	50-54
50-54	24	0.9693	23.3	-0.169	23.1	55-59	50-54	19.0	0.9693	18.5	-0.169	18.3	55-59
55-59	17	0.9482	16.1	0.292	16.4	60-64	55-59	23.1	0.9482	21.9	0.292	22.2	60-64
60-64	12	0.9175	11.0	0.671	11.7	65-69	60-64	16.4	0.9175	15.1	0.671	15.7	65-69
65-69	7	0.8627	6.0	0.345	6.4	70-74	65-69	11.7	0.8627	10.1	0.345	10.4	70-74
70-74	5	0.7767	3.9	0.100	4.0	75-79	70-74	6.4	0.7767	5.0	0.100	5.1	75-79
75-79	2	0.6624	1.3	0.111	1.4	80-84	75-79	4.0	0.6624	2.6	0.111	2.7	80-84
80-84	0	0.5164	0.0	0.076	0.1	85-89	80-84	1.4	0.5164	0.7	0.076	0.8	85-89
85+	0	0.3148	0.0	0.002	0.0	90+	85+	0.1	0.3148	0.0	0.002	0.0	90+
Total	462		498.6	47.750	546.3	Total	Total	546.3		579.3	47.750	627.1	Total
FEMALES							FEMALES						
Birth	42	0.9956	41.3	0.418	41.7	0-4	Birth	41.5	0.9956	41.3	0.418	41.7	0-4
0-4	44	0.9993	44.0	0.127	44.1	5-9	0-4	41.7	0.9993	41.7	0.127	41.8	5-9
5-9	51	0.9996	51.0	1.060	52.0	10-14	5-9	44.1	0.9996	44.1	1.060	45.1	10-14
10-14	51	0.9991	51.0	7.902	58.9	15-19	10-14	52.0	0.9991	52.0	7.902	59.9	15-19
15-19	41	0.9985	40.9	17.152	58.1	20-24	15-19	58.9	0.9985	58.8	17.152	75.9	20-24
20-24	35	0.9984	34.9	9.944	44.9	25-29	20-24	58.1	0.9984	58.0	9.944	67.9	25-29
25-29	32	0.9978	31.9	5.448	37.4	30-34	25-29	44.9	0.9978	44.8	5.448	50.2	30-34
30-34	42	0.9967	41.9	2.126	44.0	35-39	30-34	37.4	0.9967	37.3	2.126	39.4	35-39
35-39	44	0.9948	43.8	1.919	45.7	40-44	35-39	44.0	0.9948	43.8	1.919	45.7	40-44
40-44	26	0.9919	25.8	2.315	28.1	45-49	40-44	45.7	0.9919	45.3	2.315	47.6	45-49
45-49	19	0.9873	18.8	0.432	19.2	50-54	45-49	28.1	0.9873	27.7	0.432	28.2	50-54
50-54	22	0.9797	21.6	0.109	21.7	55-59	50-54	19.2	0.9797	18.8	0.109	18.9	55-59
55-59	15	0.9678	14.5	-0.397	14.1	60-64	55-59	21.7	0.9678	21.0	-0.397	20.6	60-64
60-64	10	0.9496	9.5	-0.383	9.1	65-69	60-64	14.1	0.9496	13.4	-0.383	13.0	65-69
65-69	5	0.9125	4.6	-0.187	4.4	70-74	65-69	9.1	0.9125	8.3	-0.187	8.1	70-74
70-74	5	0.8526	4.3	-0.163	4.1	75-79	70-74	4.4	0.8526	3.7	-0.163	3.6	75-79
75-79	2	0.7618	1.5	-0.030	1.5	80-84	75-79	4.1	0.7618	3.1	-0.030	3.1	80-84
80-84	1	0.6188	0.6	-0.021	0.6	85-89	80-84	1.5	0.6188	0.9	-0.021	0.9	85-89
85+	0	0.3929	0.0	-0.021	0.0	90+	85+	0.6	0.3929	0.2	-0.021	0.2	90+
Total	445		481.8	47.750	529.5	Total	Total	529.5		564.2	47.750	612.0	Total

OTHERS: Births to mothers

Age at maternity	Female start pop 2001	Female end pop 2006	Pop at risk 2001-06	fert rates	births	Age at maternity	Female start pop 2006	Female end pop 2011	pop at risk	fert rates	births
15-19	41	58.9	50	175	9	15-19	58.9	59.9	59	175	10
20-24	35	58.1	47	446	21	20-24	58.1	75.9	67	446	30
25-29	32	44.9	38	640	25	25-29	44.9	67.9	56	640	36
30-34	42	37.4	40	551	22	30-34	37.4	50.2	44	551	24
35-39	44	44.0	44	241	11	35-39	44.0	39.4	42	241	10
40-44	26	45.7	36	44	2	40-44	45.7	45.7	46	44	2
45-49	19	28.1	24	2	0	45-49	28.1	47.6	38	2	0
total				2100	88	total				2100	113

Table A.4 cont.

Illustrative ethnic projections

MIXED: Survivor computations

Start Age	Start pop 2001	Surv.Prob 2001-06	Survivors 2006	Net Mig 2001-06	Final Pop 2006	End Age	Start Age	Start pop 2006	Surv.Prob 2006-11	Survivors 2011	Net Mig 2006-11	Final Pop 2011	End Age
MALES							MALES						
Birth	28	0.9946	27.7	0.491	28.2	0-4	Birth	27.8	0.9946	27.7	0.491	28.2	0-4
0-4	31	0.9989	31.0	1.056	32.0	5-9	0-4	28.2	0.9989	28.1	1.056	29.2	5-9
5-9	37	0.9993	37.0	1.516	38.5	10-14	5-9	32.0	0.9993	32.0	1.516	33.5	10-14
10-14	39	0.9983	38.9	5.704	44.6	15-19	10-14	38.5	0.9983	38.4	5.704	44.1	15-19
15-19	30	0.9960	29.9	13.537	43.4	20-24	15-19	44.6	0.9960	44.5	13.537	58.0	20-24
20-24	24	0.9955	23.9	10.914	34.8	25-29	20-24	43.4	0.9955	43.2	10.914	54.1	25-29
25-29	21	0.9949	20.9	4.226	25.1	30-34	25-29	34.8	0.9949	34.6	4.226	38.9	30-34
30-34	30	0.9946	29.8	4.563	34.4	35-39	30-34	25.1	0.9946	25.0	4.563	29.5	35-39
35-39	33	0.9921	32.7	2.844	35.6	40-44	35-39	34.4	0.9921	34.1	2.844	37.0	40-44
40-44	18	0.9877	17.8	1.091	18.9	45-49	40-44	35.6	0.9877	35.1	1.091	36.2	45-49
45-49	13	0.9823	12.8	0.382	13.2	50-54	45-49	18.9	0.9823	18.5	0.382	18.9	50-54
50-54	17	0.9693	16.5	-0.169	16.3	55-59	50-54	13.2	0.9693	12.7	-0.169	12.6	55-59
55-59	12	0.9482	11.4	0.292	11.7	60-64	55-59	16.3	0.9482	15.5	0.292	15.8	60-64
60-64	8	0.9175	7.3	0.671	8.0	65-69	60-64	11.7	0.9175	10.7	0.671	11.4	65-69
65-69	5	0.8627	4.3	0.345	4.7	70-74	65-69	8.0	0.8627	6.9	0.345	7.3	70-74
70-74	3	0.7767	2.3	0.100	2.4	75-79	70-74	4.7	0.7767	3.6	0.100	3.7	75-79
75-79	1	0.6624	0.7	0.111	0.8	80-84	75-79	2.4	0.6624	1.6	0.111	1.7	80-84
80-84	0	0.5164	0.0	0.076	0.1	85-89	80-84	0.8	0.5164	0.4	0.076	0.5	85-89
85+	0	0.3148	0.0	0.002	0.0	90+	85+	0.1	0.3148	0.0	0.002	0.0	90+
Total	322		344.9	47.750	392.6	Total	Total	392.6		412.9	47.750	460.6	Total
FEMALES							FEMALES						
Birth	26	0.9956	26.1	0.418	26.5	0-4	Birth	26.2	0.9956	26.1	0.418	26.5	0-4
0-4	30	0.9993	30.0	0.127	30.1	5-9	0-4	26.5	0.9993	26.5	0.127	26.6	5-9
5-9	35	0.9996	35.0	1.060	36.0	10-14	5-9	30.1	0.9996	30.1	1.060	31.2	10-14
10-14	37	0.9991	37.0	7.902	44.9	15-19	10-14	36.0	0.9991	36.0	7.902	43.9	15-19
15-19	29	0.9985	29.0	17.152	46.1	20-24	15-19	44.9	0.9985	44.8	17.152	62.0	20-24
20-24	24	0.9984	24.0	9.944	33.9	25-29	20-24	46.1	0.9984	46.0	9.944	56.0	25-29
25-29	22	0.9978	22.0	5.448	27.4	30-34	25-29	33.9	0.9978	33.8	5.448	39.3	30-34
30-34	30	0.9967	29.9	2.126	32.0	35-39	30-34	27.4	0.9967	27.3	2.126	29.4	35-39
35-39	31	0.9948	30.8	1.919	32.8	40-44	35-39	32.0	0.9948	31.9	1.919	33.8	40-44
40-44	19	0.9919	18.8	2.315	21.2	45-49	40-44	32.8	0.9919	32.5	2.315	34.8	45-49
45-49	13	0.9873	12.8	0.432	13.3	50-54	45-49	21.2	0.9873	20.9	0.432	21.3	50-54
50-54	16	0.9797	15.7	0.109	15.8	55-59	50-54	13.3	0.9797	13.0	0.109	13.1	55-59
55-59	11	0.9678	10.6	-0.397	10.2	60-64	55-59	15.8	0.9678	15.3	-0.397	14.9	60-64
60-64	7	0.9496	6.6	-0.383	6.3	65-69	60-64	10.2	0.9496	9.7	-0.383	9.3	65-69
65-69	3	0.9125	2.7	-0.187	2.6	70-74	65-69	6.3	0.9125	5.7	-0.187	5.5	70-74
70-74	3	0.8526	2.6	-0.163	2.4	75-79	70-74	2.6	0.8526	2.2	-0.163	2.0	75-79
75-79	1	0.7618	0.8	-0.030	0.7	80-84	75-79	2.4	0.7618	1.8	-0.030	1.8	80-84
80-84	1	0.6188	0.6	-0.021	0.6	85-89	80-84	0.7	0.6188	0.5	-0.021	0.4	85-89
85+	0	0.3929	0.0	-0.021	0.0	90+	85+	0.6	0.3929	0.2	-0.021	0.2	90+
Total	312		335.0	47.750	382.7	Total	Total	382.7		404.3	47.750	452.1	Total

MIXED: Births to mothers

Age at maternity	Female start pop 2001	Female end pop 2006	Pop at risk 2001-06	fert rates	births	Age at maternity	Female start pop 2006	Female end pop 2011	pop at risk	fert rates	births
15-19	29	44.9	37	142	5	15-19	44.9	43.9	44	142	6
20-24	24	46.1	35	361	13	20-24	46.1	62.0	54	361	20
25-29	22	33.9	28	518	14	25-29	33.9	56.0	45	518	23
30-34	30	27.4	29	446	13	30-34	27.4	39.3	33	446	15
35-39	31	32.0	32	195	6	35-39	32.0	29.4	31	195	6
40-44	19	32.8	26	36	1	40-44	32.8	33.8	33	36	1
45-49	13	21.2	17	2	0	45-49	21.2	34.8	28	2	0
total				1700	52	total				1700	71

Table A.4 cont.

Illustrative ethnic projections

CHILD ETHNICITY ASSIGNMENT

2001-06 — **Assumed father's ethnicity given mother's ethnicity**

Mother \ Father	WHITE	BLACK	SASIAN	OTHER	MIXED	Total
WHITE	0.85	0.10	0.03	0.02	0.01	1.00
BLACK	0.28	0.65	0.01	0.01	0.05	1.00
SASIAN	0.02	0.01	0.95	0.02	0.00	1.00
OTHER	0.20	0.01	0.01	0.75	0.02	1.00
MIXED	0.45	0.20	0.01	0.04	0.30	1.00

2001-06 — **Assumed father's ethnicity given mother's ethnicity**

Mother \ Father	WHITE	BLACK	SASIAN	Other	Mixed	Total
WHITE	0.85	0.10	0.03	0.02	0.01	1.00
BLACK	0.28	0.65	0.01	0.01	0.05	1.00
SASIAN	0.02	0.01	0.95	0.02	0.00	1.00
OTHER	0.20	0.01	0.01	0.75	0.02	1.00
MIXED	0.45	0.20	0.01	0.04	0.30	1.00

2001-06 — **Births by mother's and father's ethnicity**

Mother \ Father	WHITE	BLACK	SASIAN	OTHER	MIXED	Total
WHITE	2482	292	88	58	29	2920
BLACK	30	69	1	1	5	107
SASIAN	5	3	260	5	0	274
OTHER	18	1	1	66	2	88
MIXED	24	10	1	2	16	52

2006-11 — **Births by mother's and father's ethnicity**

Mother \ Father	WHITE	BLACK	SASIAN	Other	Mixed	Total
WHITE	2347	292	88	58	29	2761
BLACK	38	69	1	1	5	136
SASIAN	5	3	260	5	0	316
OTHER	18	1	1	66	2	113
MIXED	24	10	1	2	16	71

Probability of child's ethnicity given mother's and father's ethnicity

2001-06 — **Child's ethnicity**

Mother's Ethnicity	Father's Ethnicity	WHITE	BLACK	SASIAN	OTHER	MIXED	Births
WHITE	WHITE	1	0	0	0	0	2482
WHITE	BLACK	1	0	0	0	0	292
WHITE	SASIAN	1	0	0	0	0	88
WHITE	OTHER	1	0	0	0	0	58
WHITE	MIXED	1	0	0	0	0	29
BLACK	WHITE	0	1	0	0	0	30
BLACK	BLACK	0	1	0	0	0	69
BLACK	SASIAN	0	1	0	0	0	1
BLACK	OTHER	0	1	0	0	0	1
BLACK	MIXED	0	1	0	0	0	5
SASIAN	WHITE	0	0	1	0	0	5
SASIAN	BLACK	0	0	1	0	0	3
SASIAN	SASIAN	0	0	1	0	0	260
SASIAN	OTHER	0	0	1	0	0	5
SASIAN	MIXED	0	0	1	0	0	0
OTHER	WHITE	0	0	0	1	0	18
OTHER	BLACK	0	0	0	1	0	1
OTHER	SASIAN	0	0	0	1	0	1
OTHER	OTHER	0	0	0	1	0	66
OTHER	MIXED	0	0	0	0	1	2
MIXED	WHITE	0	0	0	0	1	24
MIXED	BLACK	0	0	0	0	1	10
MIXED	SASIAN	0	0	0	0	1	1
MIXED	OTHER	0	0	0	0	1	2
MIXED	MIXED	0	0	0	0	1	16
Totals		2950	107	274	86	54	3470
Sex	Prop						
Males	0.515	1519	55	141	44	28	1787
Females	0.485	1431	52	133	42	26	1683

2006-11 — **Child's ethnicity**

Mother's Ethnicity	Father's Ethnicity	WHITE	BLACK	SASIAN	Other	Mixed	Births
WHITE	WHITE	1	0	0	0	0	2347
WHITE	BLACK	1	0	0	0	0	292
WHITE	SASIAN	1	0	0	0	0	88
WHITE	OTHER	1	0	0	0	0	58
WHITE	MIXED	1	0	0	0	0	29
BLACK	WHITE	0	1	0	0	0	38
BLACK	BLACK	0	1	0	0	0	69
BLACK	SASIAN	0	1	0	0	0	1
BLACK	OTHER	0	1	0	0	0	1
BLACK	MIXED	0	1	0	0	0	5
SASIAN	WHITE	0	0	1	0	0	5
SASIAN	BLACK	0	0	1	0	0	3
SASIAN	SASIAN	0	0	1	0	0	260
SASIAN	OTHER	0	0	1	0	0	5
SASIAN	MIXED	0	0	1	0	0	0
OTHER	WHITE	0	0	0	1	0	18
OTHER	BLACK	0	0	0	1	0	1
OTHER	SASIAN	0	0	0	1	0	1
OTHER	OTHER	0	0	0	1	0	66
OTHER	MIXED	0	0	0	0	1	2
MIXED	WHITE	0	0	0	0	1	24
MIXED	BLACK	0	0	0	0	1	10
MIXED	SASIAN	0	0	0	0	1	1
MIXED	OTHER	0	0	0	0	1	2
MIXED	MIXED	0	0	0	0	1	16
Totals		2814	115	274	86	54	3343
Sex	Prop						
Males	0.515	1449	59	141	44	28	1721
Females	0.485	1365	56	133	42	26	1621

This is a blank page

Chapter 4

Estimating methodologies and assumptions for projections: current practice for population projections by ethnic group

Ludi Simpson

4.1 Scope and summary

A variety of data sources have supplemented the decennial census to allow estimates of the ethnic group dimension of population change between censuses. One aim of this paper is to summarise the methods currently in practical use to estimate the size of the population for each ethnic group in local authority areas and smaller areas for years since 1991. These methods are given in Sections 4.2 and 4.3.

By exposing current practice to discussion, the data sets and methods may be improved for use after the 2001 Census, and a common framework adopted for national and subnational projections. Sections 4.5 and 4.6 make proposals for such a framework, describing a model and data to populate it respectively. Section 4.4 provides a bridge, drawing conclusions from the experience described in earlier sections that are used to specify the proposals in the later sections.

The experience referred to in the earlier part of this paper, after a brief review of national, regional and local approaches, focuses on the five organisations in the United Kingdom (UK) that have completed work to estimate local area populations with an ethnic group dimension:

Bradford Council
London Research Centre[1]
Oldham Council
Rochdale Council
Centre for Research into Ethnic Relations, Warwick University (CRER)

1 *The demographic and statistical functions of the London Research Centre were taken into the Greater London Authority in 2000, to become its Data Management and Analysis Group.*

Local authorities have made more progress in this area of work than other organisations. The reason is partly due to the existence of data sets that are not collated nationally, and partly due to the political priorities of local authorities in areas in which minority ethnic groups have made up a substantial proportion of the population for many years.

The projections are accepted in these local authorities because they show strong trends, which are both plausible and evidenced. The results allow high-profile policy to be implemented. These results include:

- growth in the size and geographical spread of Black and Asian populations, with political impact, and implications for the cultural aspects of all policies;
- rapid rates of ageing of Black and Asian populations, with implications for health services, social services and community development;
- profiles of future headship rates and labour force participation which are each lower than those expected if no ethnic group dimension is included, and have implications for the planned nature of housing and jobs.

The five organisations above have each been concerned to project forward from 1991 the population in each ethnic group for local authority district areas. Bradford and CRER have also added an ethnic group dimension to population estimates for electoral wards. Bradford is so far alone in adding an ethnic group dimension to estimates for smaller areas than electoral wards – a broad South Asian / Other classification.

Enfield, Luton and Birmingham Councils all aim to do similar work in the near future. There may be other one-off attempts to update the ethnic group composition of local populations, most of which will not be within the published literature. For example, Luton Council used ethnic categorisations from the 1991 Census, a schools census and a sample survey Skills Audit, to provide an estimated composition of Luton's 1999 total population (Luton Borough Council, 2000). Manchester Council used the ethnic group composition measured by the Labour Force Survey for the Greater Manchester region to deduce change since the 1991 Census in the District and ward populations of each ethnic group (Manchester City Council, 1999).

All the work included in this paper has considered ethnic group as the dimension of interest, rather than nationality or country of birth. Both these variables are occasionally mentioned, but merely as available proxies for ethnic group.

Rather than consider each organisation's method in detail, the main effort has been devoted to a summary of the data and methods in use, which is contained in two Tables in Section 4.2. While the addition of ethnic group to the national census questions has added an impetus to analysis work for all geographical areas of Britain, some useful data series are only available locally. Examples relating to fertility and overseas immigration are given in Section 4.3. The comparison between the five organisations shows that it is possible to forecast every local area using census and other data available nationally. The data that are available only locally can, for those local areas, fill the parameters of a forecasting model with more validity and reliability.

Inevitably and rightly, the nature of the available data dictates the methods of analysis that are most appropriate. It would be wrong to assume that the traditional strategies used for population projections by Government in the UK are the best when ethnic group is a dimension of the work. The focus on particular ethnic groups has, for example, had an impact on estimation strategies – particularly the focus on South Asian groups in the Pennine towns, where the analysis of names in administrative registers has been a useful demographic tool.

This review of existing practice in local and academic spheres indicates strategies to make projections of the population by ethnic group both for countries and for selected local areas. Conclusions are given in Section 4.4, and favour a single year of age, cohort progression model with allowance for ethnic group affiliation to change across generations as well as over time. Understanding past trends in vital rates and migration is required to give weight to those trends that are most likely to continue in the future. Data that are only available locally should be used to validate – and if necessary replace – initial subnational projections based on data available nationally for all areas. Successful official projections require support from users, and

political support across party boundaries and community organisations.

The details of a forecasting model and the data to populate its parameters could bear considerable further discussion. However, the technical conclusions of Section 4.4 are clear enough to indicate an outline proposal, which is given in Sections 4.5 and 4.6.

4.2 Estimation methodologies

It is unlikely that there will ever be a single best method similar in all its details, for estimating and projecting the population of each ethnic group. Due to data availability, current methods vary greatly according to the geographical scale that they address, as shown in Table 1. Survey data are used for national and regional estimates, cohort progression from the most recent census for District areas, and administrative or vital event counts for the smallest areas.

It is questionable whether the Labour Force Survey can provide more reliable estimates of ethnic group populations than cohort progression from the previous Census. The Labour Force Survey is subject to some bias (Owen, 1993) and large sampling variability in subnational areas (Scott et al, 2001; Haskey and Scott, 2001). It will be possible to compare the two approaches with the results of the 2001 Census.

The biggest weakness of cohort progression methods is the quality of migration data for each ethnic group. This weakness becomes more important as smaller, less self-contained areas are estimated. For areas smaller than local authority districts, counts of administrative records become more reliable indicators of local population. Analysis of names on electoral and patient records is currently used, but school pupil and birth records have potential here too, as in many areas ethnic group is routinely recorded.

Through mixing and matching and squeezing methods appropriate to incomplete data of variable quality, a set of suitable estimates are achieved by local practitioners. The state of the art of ethnic group population statistics is more similar to small area demography and to third world demography than to current UK government demographic work.

Table 1

Methods used to estimate populations by ethnic group since 1991: national, regional, district and smaller areas

Geographical scale	Methods	Advantages	Disadvantages
National and regional	• Direct estimation from Labour Force Survey. *ONS, Greater Manchester, Birmingham.*	• Nationally consistent. • Characteristics beyond age and sex. • Sample improving with LFS boost in city areas.	• Known bias within Labour Force Survey which under-estimates minority groups. • Small sample for sub-regional areas even with boost.
District	• Direct from a local census. *Oldham.*	• Characteristics beyond age and sex.	• Bias to be expected as for surveys.
	• Cohort component progression since last census. *Bradford, London, Oldham, Rochdale, CRER*	• Fine age-sex groups. • Anchored in relatively high-quality census data. • Same method extends to forecasting. • Makes use of existing models that do not have an ethnic group dimension.	• Heavy requirement to estimate each component of change. • Migration within the UK is the least robust component.
Smaller areas	• Direct from a local census. *Oldham.*	• Integrated with larger enquiry.	• Bias to be expected as above.
	• Cohort progression since last census. *Bradford, CRER.*	• Fine age-sex groups. • Anchored in census data.	• Migration not measured, in spite of its importance.
	• Scaled counts from electoral and patient admin. records (synthetic estimation, 'apportionment'). *Bradford.*	• Provides postcode building bricks for neighbourhood statistics. • Patient data give fine age-sex groups.	• Ethnic group not directly coded on these data sets, but derived from name-analysis.

Table 2 summarises the methods and data currently used in practice in the UK to estimate and project populations by ethnic group. Table 2 also includes the attempts of five organisations to estimate or project for local authority districts and smaller areas. The three authorities that span the Pennines are grouped as they have collaborated in their work and use very similar approaches and data sources. The Table concludes with references to published accounts of each method.

Table 2 is intended for reference, but two general comparisons between the approaches are worth making in advance of the issues raised later in the paper.

First, all five organisations use standard software as a platform on which programmed routines perform relatively standard implementations of cohort progression forecasting models. There is nothing particularly original about the models, except for the inclusion of a propensity to change ethnic groups between mother and child within the Bradford/Rochdale model.

Second, there is a strong gradient in the extent that the organisation uses data that have been gathered especially for the purpose of estimating populations by ethnic group. The CRER estimates of 1996 population use no such data, relying solely on data from the 1991 Census. They estimate fertility rates from it and progress the census age structure forward; migration and mortality are then assumed equal for each ethnic group when scaling the initial results to the government population estimates for each District. The same approach could be taken to project forward any number of years, and for all local areas.

In contrast, the London Research Centre, and even more so the Pennine Authorities, gather other data in order to estimate the variation in fertility, mortality and migration between each ethnic group. It is no accident that the extent to which local data are gathered varies inversely with the number of authorities projected: most local data is used by organisations projecting a single local authority area, some by the organisation projecting 33 London authority areas, and none at all by the organisation concerned with all areas within Britain.

Table 2

Estimation methodologies for subnational ethnic group population estimates and projections

	Pennine metropolitan councils: Bradford, Rochdale, Oldham	London Research Centre For each of 33 London Boroughs	CRER, Warwick University For each local authority District and ward in Britain.
(a) POPULATION: COHORT COMPONENT ESTIMATES AND PROJECTIONS, MALE AND FEMALE SEPARATELY			
Software	*B&R:* POPGROUP Excel-based, commercially available. *O:* In-house, Excel-based.	In-house, Excel-based.	In-house, SAS.
History	*B*: Approx each three years since 1979, latest 1999-based. *R:* One round, 1999-based, commissioned from Bradford Council. *O:* One round, 1996-based.	Three rounds published 1995, 1997, 1999.	1991 rolled forward to 1996 only.
Characteristics	B,R,O: Single year of age. Six ethnic groups (As 1991 Census with 'African-Caribbean' and 'Chinese and other' each amalgamating three census groups).	Single years of age. Ten ethnic groups. Each Borough estimated separately, the sum of ethnic groups constrained to an independent total projected for that Borough.	Five-year age groups. Four ethnic groups. Migration not explicit, but applied by constraining the sum of ethnic groups to the ONS mid-1996 District population estimate.

cont.

Table 2 cont.

Estimation methodologies for subnational ethnic group population estimates and projections

	Pennine metropolitan councils: Bradford, Rochdale, Oldham	London Research Centre For each of 33 London Boroughs	CRER, Warwick University For each local authority District and ward in Britain.
(a) POPULATION: COHORT COMPONENT ESTIMATES AND PROJECTIONS, MALE AND FEMALE SEPARATELY			
1991 base population. In each case: (a) 1991 Census is used (table L06 with five-year age groups). (b) Allocation of census undercount was specific to each age & sex, and controlled to the ONS estimates. (c) Composition of student transfer assumed to be general, or undocumented.	B,R,O: single year of age imposed from all-groups population. B: undercount allocated (by GMAP Ltd) in each ward in proportion to imputed households. O: undercount allocated (by ETHPOP Univ.Mcr routine) in each ward as per Estimating with Confidence. R: undercount allocated (by SOCPOP Univ.Mcr routine) in each ward, with specific differential for each ethnic group.	Single year of age taken from specially commissioned Table. Undercount allocated equally to each group.	Undercount allocated in each ward as per Estimating with Confidence (as Oldham)
Birth events	Maternity records tabulated for: *B:* since 1978 *O:* since 1991 *R:* since 1994 In each case, coped with incomplete data and with non-Census categories (eg 'Mediterranean', 'Mother West Indian/ Father White').	Not measured	Not measured
Fertility rates	*B,O,R:* ASFRs and TFR estimated annually from birth events, trends observed. All show high but decreasing Pakistani and Bangladeshi TFR, and very low Indian and Black TFR. *B:* Some work on differences between UK-born and other mothers of each S.Asian ethnic group (showing lower fertility of UK-born).	TFR derived from the Census ratio 0yr/(15–49yr) (for smallest groups, regional ratios used). Age pattern copied from the regional ASFRs of a relevant Country of Birth of mother (eg UK for White).	Census ratio (0–4yrs)/(15–45yrs) applied directly to the 15–45 yrs in 1991
Mixed parentage	B,O,R: As LRC. *B:* Since 1996, both mother and child's ethnic group recorded, giving a matrix of propensities that a child of a mum of group Y will be in group X. Such a matrix is accepted in POPGROUP software, but not used as yet.	Mixed parentage implicitly taken into account as follows: children of group X born to mothers of group Y are included in the fertility rate for group X.	As LRC
Sex ratio	*B,O,R:* small differences varying over time noted but not built into model.	Small differences between Country of Birth of mother, built into model for relevant ethnic group.	Assumed no difference.
Death events	Not measured except for 1998–92 study (see 'mortality rates').	Not measured	Not measured

cont.

Table 2 cont.

Estimation methodologies for subnational ethnic group population estimates and projections

	Pennine metropolitan councils: Bradford, Rochdale, Oldham	London Research Centre For each of 33 London Boroughs	CRER, Warwick University For each local authority District and ward in Britain.
(a) POPULATION: COHORT COMPONENT ESTIMATES AND PROJECTIONS, MALE AND FEMALE SEPARATELY			
Mortality rates	B,R,O: Use a small sample study in Bradford 1988–92 based on Country of Birth and name origin of deceased, showing SMRs for age 35–74, England=100, as follows: Indian M 147 F 116 Pakistani M 102 F 84 Bangladeshi M 152 F 109 All others M 115 F 109 B,R,O: Infant mortality – Pakistani, Bangladeshi and Black groups have 1.6 times the infant mortality of other groups, from ONS and Bradford evidence.`	Assumed equal. National evidence is contradictory for all-cause mortality; differentials for cause-specific mortality by country of birth tend to cancel each other out.	Assumed equal.
Migration events within the UK	Not measured	Not measured	Not measured
Migration rates within the UK All calculate out- and in-migration rates or propensities, using the LAD population as the denominator.	*B,R,O:* Age-specific rates from the NHSCR for period around 1991 act as starting point and constraint. *B:* South Asian groups' out- rates decreased and other groups' out- rates increased, to reflect the change in electorate 1991–99 as evidenced by name analysis. Census evidence considered not to apply (incomplete and atypical). *O:* NHSCR rates scaled to meet the census evidence from SMS and LBS tables.	Special commissioned 1991 Census table gives age-sex-group specific in- and out- rates.	Assumed equal at each age and sex (net impact of all migration is derived implicitly by constraint to ONS mid-1996 population estimate).
Migration events with overseas	*B:* Home Office 'Port Health notifications' of first address of immigrants, used for age-sex counts for each of Pakistani, Bangladeshi, and Indian. 1991 Census SMS Table 5 by ethnic group and by Country of Birth, examined but not used. The IPS annually 1975–98: Country of Origin/ Destination, by age groups 0–14, 15–59, 60+, by sex, for Bradford and region; examined, gives long-term trends and composition, but not used in estimates. *O,R:* As Bradford, with exception of IPS which was not available (sample size). *R:* All flows other than Bangladeshi, Pakistani and Indian immigration, derived by subtraction from the projected ONS total. *B,O:* These other flows assumed to have zero net impact.	Asylum seekers and visitor switchers. The ONS figure for all London each year is distributed using LRC's own study of their composition, to Boroughs and with detail of age-sex-ethnic group. This study was based on information from social services, schools, and housing services, and the country of origin given with Home Office statistics of asylum seekers and visitor switchers. These are used to supplement immigration events estimated by applying rates as below.	Not measured

cont.

Table 2 cont.

Estimation methodologies for subnational ethnic group population estimates and projections

	Pennine metropolitan councils: Bradford, Rochdale, Oldham	London Research Centre For each of 33 London Boroughs	CRER, Warwick University For each local authority District and ward in Britain.
(a) POPULATION: COHORT COMPONENT ESTIMATES AND PROJECTIONS, MALE AND FEMALE SEPARATELY			
Migration rates with overseas	Estimated from immigration events for Indian, Pakistani, Bangladeshi as above.	The special 1991 Census table giving UK migration also provides rates of immigration by age and sex. Emigration estimated from the IPS using an average of three years.	Assumed equal (as above for migration rates within the UK).
Boundary changes	Not considered	Considers ethnic composition of areas affected by boundary change since 1991.	Subsumed within the constraint to ONS mid-year estimate
(b) HOUSEHOLD ESTIMATES AND PROJECTIONS			
Households	*B:* From names of electors, analysis of changing number of households and frequency of adults sharing	Not measured	–
Headship rates	*B:* Local rates estimated by five-year age, sex, ethnic group and consistent with those used by DETR for 1991, by Iterative Proportional Fitting from Census SARs, LBS, and special commissioned tables. Population in households estimated similarly. Nationally, the relatively low headship rates among South Asian groups are stable both over time (from Longitudinal Study 1981 to 1991) and between generations (from 1991 Census SAR). *O:* Local rates estimated for 1991 using special census tables, as Bradford.	Household size derived from Census, using table of ethnic group by head of household.	–
(c) LABOUR FORCE ESTIMATES AND FORECASTS			
Labour force	Labour Force Survey provides only White and Other, due to small local sample.	–	–
Economic activity rates	*B,R:* Local rates estimated by five-year age group, sex, ethnic group and consistent with the 1991 Census, by Iterative Proportional Fitting from Census SARs, LBS, and national topic volume. *B:* Nationally, the very low activity rates for Pakistani and Bangladeshi women after age 29 have been stable over time.	–	–

cont.

Table 2 cont.

Estimation methodologies for subnational ethnic group population estimates and projections

	Pennine metropolitan councils: Bradford, Rochdale, Oldham	London Research Centre For each of 33 London Boroughs	CRER, Warwick University For each local authority District and ward in Britain.
(d) SMALLER AREA ESTIMATES AND FORECASTS			
Wards and EDs	*B:* Estimates for 1996 and 2000 in Census EDs approximated by whole postcodes. Two methods used, the estimate being the mean of the two: (a) cohort progression from a 1991 population based on the census, (b) triangulation of postcoded patient and electorate counts, and District estimates, using a dichotomous ethnic group from name analysis of administrative records (South Asian/Other) which could be extended.	LRC receives from ONS live births by country of birth of mum, by ward (Table VS4A).	Cohort progression from 1991 to 1996 in England and Wales, constrained to District estimates as above, each ward in England and Wales. Unpublished.
(e) PUBLICATIONS AND DOCUMENTATION			
Publications and documentation	*B:* Simpson, 1997, 1998. Andelin and Simpson, 2002. Bradford Council, 1996, 1997, 2000. Simpson and Close, 1998. Harding et al, 1999. *O:*Oldham Metroplitan Borough Council, 1995, 1997 *R:* Bradford Council, 1999	Storkey, 2000	Owen, 1999

Notes to Table 2

'Not measured': in these cases, each group's events are estimated – usually from rates applied to a previous population, and scaled to meet the events measured for the all-groups population.

'Regional': in these cases, the LRC applied data for Inner London and Outer London to Boroughs where data was sparse.

4.3 Examples of local data

Four examples are given to show the variety of data that can be brought to bear on population projections with an ethnic group dimension. Some may yet evolve into a national data source. At present they are candidates to verify locally other estimates that do have national coverage but may be less accurate.

The purpose of the section is to demonstrate the value of these data sets. It does not pretend to review in detail each one of them; further background and analysis can be found in Simpson (2000) and Storkey (1999a). The four examples refer to fertility and immigration.

4.3.1 Births and fertility rates

Following work by the former Sheffield City Polytechnic (Drew and Simpson, 1981), a time series of birth counts in Bradford with detail of ethnic group, age of mother, parity and sex has been maintained since 1971. Since 1978 this has been derived from a database of individual live birth records copied from Health Authority maternity and child health administrative systems. Categories for ethnic group of the child, and administrative arrangements for the database have changed over time, but careful estimation, in conjunction with Census-based population data, has allowed the age-specific fertility rates for ethnic groups, and the Total Period Fertility Rates (TFRs) (shown in Figure 1) to be estimated.

The TFR of over ten in the early 1970s does not indicate families of that size. The extreme TFR for that period reflects the nature of immigration to Bradford at that time, comprising mainly women from Pakistan and Bangladesh, who were beginning or continuing to have children, having been separated from their husbands.

The relatively high Pakistani and Bangladeshi fertility rates are therefore partly a reflection of immigration status, and not of ethnic group. The Bradford forecast therefore expects a continuing fall in fertility rates for those groups, as immigrant spouses are a decreasing proportion of adult women.

Indian and African-Caribbean fertility in Bradford is now lower than White fertility, as in Rochdale and Oldham, and a national study based on the Longitudinal Study (Storkey, 1999b). This is partly due to the fertility being calculated as a ratio of children recorded as ethnic group X divided by the population of ethnic group X. Populations that are reduced by mixing with other ethnic groups have reduced measured fertility, as their children are not all included in the calculation. Conversely, these children contribute to a relatively high measured fertility in the 'Other groups' category that they are included within.

4.3.2 Ethnic group of child and mother

Since 1995, the same database of births has recorded ethnic group of both mother and child for births within the city area of the Bradford District. Three years of this data is presented in Figure 2.

'Mixed' is not an ethnic group category in the current Health Authority classification that is based on the 1991 Census. Of the 510 births recorded as Other, 320 have mothers of a specific origin and so may be assumed to be of mixed parentage. Of note is that a quarter of Caribbean mothers have children recorded as Other ethnic group. The recording is made by the maternity service as part of the hospital record, with the involvement of the mother.

However it is not the norm that the 'Other' category is being used as a substitute for 'Mixed'. There are 191 cases in the off-diagonal entries of the above Table, where the child's ethnic group is different from the mother's ethnic group and is not 'Other'. These births, too, are likely to be of mixed parentage, where the child has been allocated the father's or another specific ethnic group. The births where the mother or the child's ethnic group is 'Not given' have been excluded from these figures.

It is in recognition of these characteristics, that recent software used in Bradford allows a matrix showing the propensity of births to mothers of one ethnic group to belong to the same ethnic group or each of the other ethnic groups (Andelin and Simpson, 2002).

4.3.3 Immigration: Home Office Port Health Notifications

Since 1972, Bradford Council has maintained counts of the number of Port Health Notifications sent by Home Office officials from each port and airport to the Health Authority containing the first address of each immigrant undergoing a routine health check. It is thought to include all immigrants from New Commonwealth countries and many others besides. The notification contains intended first address, age, sex, nationality, origin, and a 'refugee/student/short stay/long stay' status. The counts by Bradford Council exclude students and short stay immigrants. Port Health Notifications have also been monitored by Calderdale, Oldham and Rochdale Councils from their respective Health Authorities, but other Councils have not been able to trace them, and they are not collated centrally by the Home Office.

Comparisons with census and International Passenger Survey data, and analyses of age-sex structure, suggest that the Notifications provide a reasonably accurate picture of immigration.

Figure 3 shows the time series of Port Health Notifications, divided into children under 15, and males and female adults (aged 15 or over). It has accorded well with our understanding of the main reasons for changes in the levels of immigration, marked on the chart.

Figure 1

Total Fertility Rate in Bradford District

From 1999, the figures are future trends assumed for the 1999-based forecast

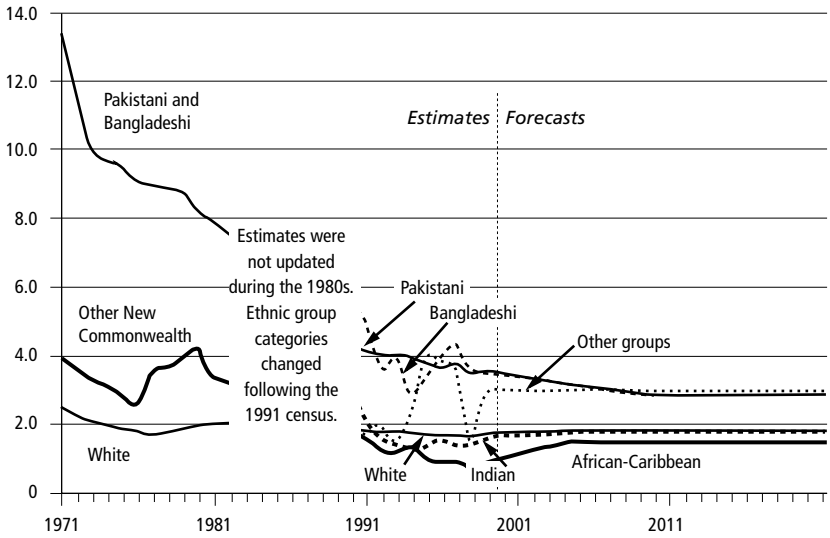


Figure 2

Extract from City of Bradford Metropolitan Council - Births Statistics System - Ethnic group of mother and child

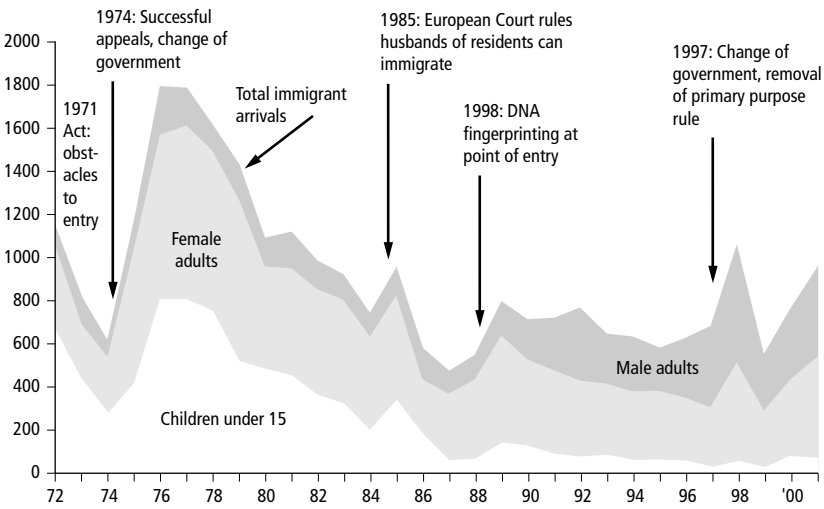
Period 01/07/1995 to 30/06/1998
Bradford area and All births

	Ethnic group of child										
Ethnic group of mother	White	Caribbean	African	Black-Other	Indian	Pakistani	Bangladeshi	Chinese	Other	Not given	Total
White	9,042	4			3	35	4	1	257	6	9,352
Caribbean	6	70		1					26		103
African			26	1		1			2		30
Black-other	1	1		31	1	1			2		37
Indian	1		1	1	517	5	1		14		540
Pakistani	20		2	5	23	5,358	25		18	2	5,453
Bangladeshi	5			1	1	15	386		1		409
Chinese	1							20			21
Other	20		1	1	2	9	1		187		221
Not given	143		1		10	73	6	1	3		237
Total	9,239	75	31	41	557	5,497	423	22	510	8	16,403

The number of successful applications to immigrate is not as unstable as it may at first seem from Figure 3. Each trough and peak is associated with administrative changes in the UK government's regulations regarding immigration, which cause backlogs of applications and then recoveries to deal with the backlogs after regulations are relaxed. This suggests that the demand for immigration to Bradford has been steadier than the time series of successful applications suggests, and that this demand is not highly affected by the changing regulations. Thus Bradford Council uses the annual data for updating estimates of population, and uses the longer-term average for forecasting the likely future level of immigration.

Figure 3

Port Health Notifications of immigrants to Bradford District



4.3.4 Immigration: asylum seekers

The number of those seeking asylum who become new UK residents has been around 25,000 each year during the late 1990s, according to government population estimates. Their geographical distribution is not known at all precisely. Since they are likely to be concentrated in areas of diverse ethnic group composition, but themselves not reflect the ethnic group composition of their host community, the London Research Centre has documented fifteen sources of information that can better locate the distribution of residents seeking asylum in Britain (Storkey, 1999a).

The London Research Centre undertook a comparison of the most relevant estimates for refugee distribution within London, which is

reproduced in Figure 4. The column headings refer to data sources as follows:

Terminal 3: New arrivals referred to the Health Control Unit at Heathrow Terminal 3 and recorded as Asylum Seekers, April 1997–March 1998. This is a computerised dataset and drawn from the same administrative immigration procedures described in Section 4.3.3 above.

Social Services: Case load of Asylum Seekers supported under the National Assistance Act and the Children Act, January 1997–February 1999.

BABIE: Bed and Breakfast Information Exchange. Homeless asylum seekers 1990–1997.

Refugee Council children: Estimates of refugee school children collated by the Refugee Council from a variety of survey and other sources, 1996–1997.

ONS MYE: Asylum seekers and visitor switchers included as new residents in the Office for National Statistics Mid-year estimate, 1996–1997. This reflects the country of birth of residents from the 1991 Census, for those regions of the world from which the Home Office reports asylum seekers and visitor switchers.

The authors of the Table reproduced in Figure 4 discuss each data source and conclude:

'Although there are some anomalies between the different methods, and each has its own weaknesses, there is a reasonable degree of consistency. The Terminal 3 data is likely to be the least representative, yet the results are surprisingly consistent. The ONS data could be said to be an indicator of more settled communities, whereas the BABIE data will be influenced by the amount of temporary accommodation in each area. … these distribution assumptions have a fairly small effect compared to the uncertainty about how many refugees there are in the country' (Aldous et al, 1999, page 23).

The London Research Centre for its population projections by ethnic group, has used the mean of the estimates listed in Figure 4, excluding that headed 'Terminal 3'.

Figure 4

A comparison of the different estimates of the percentage of the total refugee community in London's Health Authorities

Reproduced from Table 2.8 in 'Refugee health in London' (Aldous et al, 1999).

	Estimate based on data from:				
Health Authority	**Terminal 3**	**Social Services**	**BABIE**	**Refugee Council children**	**ONS MYE data**
Barking & Havering	1%	2%	3%	1%	1%
Barnet	3%	3%	3%	5%	7%
Bexley & Greenwich	11%	4%	2%	5%	2%
Brent & Harrow	9%	5%	7%	9%	9%
Bromley	1%	1%	1%	0%	2%
Camden & Islington	5%	9%	14%	6%	7%
Croydon	2%	3%	4%	2%	3%
Ealing, Hammersmith & Hounslow	14%	9%	10%	11%	13%
East London & The City	17%	16%	11%	22%	11%
Enfield & Haringey	3%	11%	8%	13%	7%
Hillingdon	11%	2%	2%	1%	2%
Kensington & Chelsea and Westminster	7%	10%	13%	5%	12%
Kingston & Richmond	0%	2%	2%	1%	3%
Lambeth, Southwark & Lewisham	9%	15%	12%	9%	10%
Merton, Sutton & Wandsworth	4%	4%	4%	6%	7%
Redbridge & Waltham Forest	2%	5%	3%	5%	4%

4.4 Creating assumptions for projections: issues

4.4.1 A forecast rather than a projection

The purposes outlined by ONS for projections with an ethnic group dimension involve the setting of targets for key government policies in relation to social exclusion (Haskey, 2000). Population estimates and projections would guide the future employment, housing and take-up of services expected from each community if equal opportunities were to be achieved. The motivation of local authority work outlined in Section 4.2 has also been to guide policy through knowledge of the most likely future demographic composition of local areas.

Work preparatory to projections reviews past changes in ethnic group fertility, immigration, economic activity, headship rates and other components of demographic and related change. Only trends that are likely to continue are built into the projections. While a range of projections might be attempted to suggest alternative

plausible scenarios, the users of ethnic group projections demand that the most likely scenarios are indicated, on the basis of sociological and demographic evidence.

There is little point developing projection models of demographic change, unless it is to allow information to be included that will best show what is predictable about the future.

4.4.2 The convergence of ethnic group vital rates cannot be assumed

There is clear measured diversity among the ethnic groups as regards vital rates and key rates of prevalence for projections.

- Fertility: high for Pakistani and Bangladeshi groups, low for the Caribbean group.
- Mortality: variation between groups associated with specific causes of death.
- Migration within Britain: relatively high migration between existing settlements of ethnic groups.
- Migration overseas: net immigration for ethnic groups other than White.
- Economic activity: low activity for Pakistani and Bangladeshi men; very low activity for Pakistani and Bangladeshi women after age 29.
- Headship rates: low headship rates (or equivalently larger household size) for the Indian, Pakistani and Bangladeshi groups, indicating sharing between families of the same generation as well as across generations.

When making a projection of future demographic change, some commentators suggest that the impact of schooling on the children of immigrants, and other cultural and economic pressures, will lead to a convergence of these vital rates and social indicators.

In some cases, the reasons for the ethnic group diversity are unknown or the differences are considered unpredictable – such as is the case for migration patterns between London Boroughs and overall all-cause mortality rates. In these cases a reduction of the diversity has been used as a 'conservative' option to best represent the uncertainty.

However, in cases where evidence exists, it does not always point to convergence of these rates. Overseas migration is a continuing phenomenon for some of Britain's ethnic groups that are truly international communities with marriage markets spanning continents. This has been shown with data for Bradford and the three South Asian ethnic groups (Simpson, 1997). Low headship rates for each South Asian group have persisted in Britain over time and between generations (Bradford Council, 1997), and the low Pakistani and Bangladeshi age-specific economic activity rates have persisted in Britain for over a decade (Bradford Council, 1996).

For mortality, there is evidence that first generation migrants are relatively healthy with lower mortality than later generations, leading to a widening of mortality differentials between ethnic groups (see Seeromanie Harding's chapter in this volume).

Only in the case of fertility rates have the high Pakistani and Bangladeshi rates been clearly decreasing over time. Fertility rates are lower for mothers born in Britain (Bradford Council, 2000; Blackwell, 2000, similarly shows higher childlessness among those born in Britain). The reduction over time, however, affects mothers irrespective of place of birth (Bradford Council, 2000). All the forecasts discussed in Table 2 assume a reduction of Pakistani and Bangladeshi fertility rates in the future. They also assume an equalisation of fertility rates for other ethnic groups, though this assumption is not backed by evidence.

It would be quite wrong to adopt a false stereotype either of convergence of cultural norms and aspirations, or of unrelated experiences, which do not accord with reality. Careful measurement of demographic rates over time and between generations is a high priority for ethnic group projections, in order to inform assumptions about the future with evidence from the past.

4.4.3 The current age structure determines the future: cohort progression favoured

It is a rule of population forecasting that the future age-sex structure is more predictable than the total population. The future age of each individual surviving ten years will be greater by ten years.

The age-sex structure of each ethnic group is known from the most recent census. The groups other than White have a young age structure that indicates future growth, as births will outnumber deaths for many years to come. The Bangladeshi immigrant families, now re-united in Britain, have more older men than women, reflecting the large difference in age at marriage within Bangladesh. This sex imbalance will persist into the future for the British Bangladeshi elderly population.

It is therefore important that projections are made using a cohort progression methodology, and that they use a single year of age population structure for their base.

4.4.4 Modelling mixed parentage and the changing definition of ethnic group

An individual's ethnic group, as officially recorded, changes with all three dimensions of demographic models – over time, with age, and between cohorts. Ethnic group is dependent on official categorisation, on cultural and political attitudes to recording one's race or ethnicity, and on the nature of the person recording ethnic group. The latter, for example, changes during a lifetime from midwife to parent to self-assessment. These problems of definition for demographic work are reviewed in Simpson, 1997.

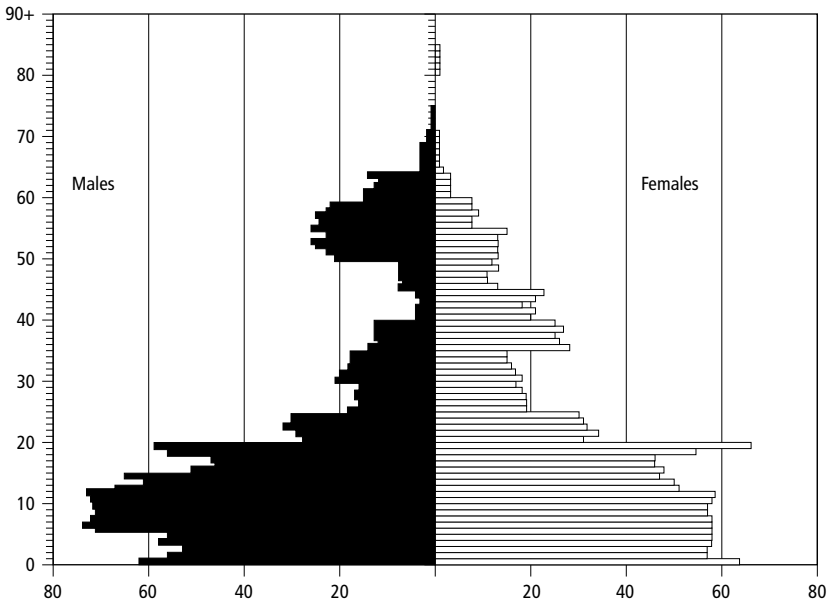
In a rather different way, ethnic group changes within a family for children of mixed parentage.

All these changes can be modelled in population projections using a simple square matrix of propensities for changing ethnic group from A to B. In the POPGROUP forecasting software, such a matrix is applied to the births of women of each ethnic group. It could equally be applied in an age-specific way, to existing persons as they age from one forecast period to the next.

A practical alternative for mixed parentage births used by current practitioners of Table 2, though less theoretically satisfying, is to compute fertility rates as ratios of births of ethnic group X divided by mothers of ethnic group X, where the births include those to mothers of other ethnic groups. This tends to depress the 'fertility rates' measured for the groups that inter-parent, and inflate those for the group(s) that contain mixed births.

Figure 5

Bangladeshi estimated population, mid-1991, Bradford



Another alternative is to define ethnic group as 'adult ethnic group', ascribing to children the ethnic group that they are likely to adopt when adult, based on past evidence of the changes in ethnic group on transitions from birth to childhood, and from childhood to self-assessment. Such evidence would need to be gathered for contemporary Britain. This discussion also leads towards Berthoud's (1998) conclusion that for demographic analysis one should focus for every individual on their parents' family origin(s), but these have not been the subject of questions in the UK Census or the Labour Force Survey.

4.4.5 A standard model and data for estimation and projection, with local improvements

A cohort progression model, with separate components for births, deaths and migration, should be adopted as standard for projections of the ethnic group composition of national and local populations.

Data from the Labour Force Survey and administrative data available nationally or locally are able to improve the estimates of population and demographic rates in subsequent years after a national census.

Development of population statistics for ethnic groups should therefore include methods of incorporating these extra data into the early years of a projection. The implications for change, in particular for migration, can then inform the assumptions for future demographic rates. Analysis of these data will involve:

- calibration of non-Census data against the 2001 Census;
- use of estimates from national surveys to improve draft projections;
- use of local data for components of change to improve nationally-made draft projections.

4.4.6 The use of an all-groups projection to constrain the sum of group projections

One claim against constraining ethnic group projections to a total projected independently, is that the projection of a total population by summing separate ethnic group populations will be more accurate than when projecting the population as one whole. The fertility, mortality, and migration characteristics of the growing parts of the population will be properly represented in the future. Bradford, Rochdale and Oldham have forecast their total population in this way, adding up the separate forecasts of each ethnic group.

However, in all local authority projections so far including those from Bradford, Rochdale and Oldham, the assumptions about future migration within the UK have been set for the whole population as a constraint, before the ethnic group composition of these flows are forecast. When estimating the rates of fertility, mortality and migration for each group – the data whose quality is rarely as good as that for the population as a whole – the projections are again constrained to be consistent with the all-groups data. Assumptions about the future of each ethnic group are considerably less reliable (in the sense of less likely to be close to the future values, as that is what one is interested in) than those for the population as a whole. And, thus again, a constraint to an all-groups projection could be favoured as by LRC and CRER.

Thus the experience is mixed and does not clearly favour one approach over another. The quality of available data should decide this issue.

4.4.7 Single region model, repeated for each subnational area.

All the models discussed in Section 4.2 are for single regions. Even when results are published for more than one area, as in the work of London Research Centre and CRER at Warwick, the projection exercise is essentially repeated for each area. In contrast, ONS's current practice is to make subnational projections that are multi-regional in their treatment of migration; the impact of migration on any one area is affected by the size of the population in other areas, through a matrix of migration propensities.

This choice is a key one for the proposed projection model. Do the census data that would be used to construct a group-specific multi-regional migration propensities matrix provide stability over the projection period? This question has to be answered relatively, compared with the alternative single region model, where the propensity to migrate into or out of each district is calculated separately, and as a proportion of the currently resident population of that group.

4.4.8 Who decides the model, its parameters, and publication?

The great uncertainty about the future makes it important that users accept the basis on which published scenarios have been reached. In forecasts not involving an ethnic group dimension, at national and local level, there is usually a discussion involving users about evidence for alternative scenarios and the reasons for a particular choice. For some projections, these discussions occur regularly and formally, as with the ONS long-term subnational population projections. In other cases, the discussion is embodied in a historical agreement on the principles by which assumptions should be made. However it is achieved, the act of involving users in formulating the assumptions on which a forecast is based provides an important basis for the acceptance of – and support for – the final results.

Equally importantly, it is imperative to gain widespread support for government publication of projections that will forecast the growth of multi-racial Britain. Such a publication will be referred to not only when planning services but in policy debates on race and immigration, which have a volatile history in Britain.

In my opinion, two types of support are important to gain. First, community representatives should be involved in reviewing the assumptions made to set the model's parameters, perhaps including the Commission for Racial Equality. Second, political support is needed, across party lines, and that support made public, before the projection results are published.

4.5 A model for projecting populations by ethnic group

This section proposes a specific model. The limits that any model imposes on the analytical outputs should be discussed. The proposal is designed to meet the expressed needs for the projections (for example in Haskey, 2000), and to build on existing experience of population projections by ethnic group, as reviewed in Section 4.4 above.

The proposed model takes a relatively simple approach. It is motivated by the detail required by useful outputs; includes making the best use of available data, and adds detail only where it will have an appreciable improvement in the robustness of the resulting outputs.

There is a strong tendency for projection models to become complicated in order to replicate real-life processes of migration and family formation. Such complications are necessary to investigate the determinants of changing vital rates, but are impractical for regular official projections because the task of projecting the determinants themselves is seldom possible.

All aspects of the proposed model have already been tested in demographic models. Putting together the model in the way suggested for projecting the population by ethnic group will be a new development in the UK.

It is proposed that the same model be applied to national and subnational data. For subnational projections, collapsing some categories may be desirable when setting up the model. In both national and subnational projections, collapsing categories before publication would ensure greater robustness of the results (see ethnic group, age, and constraints below).

Geography

- (a) National (b) Selected Districts, plus regional remainders.

If Local Authority Districts were selected with minority ethnic proportion of their population larger than 10 per cent, as measured in the latest Census, 45 Districts would be chosen in 1991 (including an allowance for undercount), and about seventy after the 2001 Census (with all but eight London Boroughs being included).

Ethnic groups

- National: ten groups from 1991 Census. 14 in 2001, with Indian subdivided by Sikh, Hindu, Muslim.

- For the subnational projections, the model or published results from it might show only groups that have more than 1,000 residents for at least one year during the forecast.

Projection period

- 12 months (mid-year to mid-year).

Time Horizon

- From the year of the most recent data, a maximum of 20 years.

- Scenarios to show the impact of feasible demographic rates that imply high or low growth may be required to go further.

Sex

- Male and female, separately.

Age

- Single years of age.

Published results would use age groups only – which would be suitable for applications.

Demographic projection approach

- Cohort progression, with components for births, deaths, and flows of in- and out-migration with other parts of the UK and with overseas.

A propensity to change ethnic group

- An age-specific matrix to show the propensity for a person of ethnic group X at the

beginning of a projection period to end it in group Y. For each age, the matrix is square; each row sums to 1, since no person is lost from the population in this process.

For the 'newborn', those born during the projection period, this matrix allows the modelling of mixed parentage.

For older ages, the matrix allows for change in affiliation and in recording procedures.

Where data are not available, a diagonal matrix of ones maintains constant ethnic group without change.

Flexible input options

Allow for lack of detail in data and assumptions as follows.

- Allow each component of births, deaths, migration flows, to be input either as a count of events, with or without age-sex detail, or as a rate, again with or without age-sex detail.
- Allow separate inputs for each ethnic group, or for aggregates of groups, or for all groups combined.
- Allow input of net-migration or gross flows.

The detail not provided in the input would be inferred by use of default rates acquired from national analysis.

Constraints to an independent projection

Allow constraints as follows, to ensure consistency between geographical scales. In each case the model will revise previous migration estimates to meet the constraint:

- to national projection of all-groups;
- to subnational projection of all-groups;
- to national projection of each group separately;
- to subnational projection of each group separately.

The constraints to the all-group projections will force, if desired, consistency with existing government population projections without an ethnic group dimension.

The constraints ensuring agreement with existing projections will allow Labour Force Survey estimates for a post-census year to be incorporated. In these cases, the revised post-census migration will give information to improve assumptions of future migration.

The constraint ensuring agreement with the total numbers of births, deaths, and each migration stream would be possible using the flexible inputs options above.

Single-region approach

- Single regions

Projection repeated for each subnational area, with optional constraints as outlined above.

Projections of the number of households and the labour force

- If incorporated at the same time as the population projections, these will give considerable added value.

The analyses they require of Census and Labour Force Survey data will be more efficiently specified and obtained at the same time as the demographic analyses required by the core aim of the projections working group.

For these extra projections, the general approach should follow that for the official projections with no ethnic group dimension. That is, projected headship rates and projected economic activity rates should be used with the population projection. For the household projection, the number of residents not living in households must also be projected and deducted from the projection of the whole population.

4.6 Analysis of the 2001 Census and other data sets to estimate parameters for a projection model

To provide the projection model's parameters, estimates are required for the base population and for age and sex specific vital rates for each ethnic group.

In order to discern how rates are likely to change over the projection period, the most up-to-date estimate of each rate must be supplemented,

wherever possible, by estimates of change over time, and between generations (meaning here between those born outside the UK and others; see section 4.4.2).

Good practice would involve testing and validating the models and their parameters using 1991 as a base year to make projections to 2001, before the 2001 Census results are available for analysis.

The following list of analyses of the 1991 Census, the 2001 Census, and the Labour Force Survey would provide a data set with which to make projection assumptions. It is minimal for the accurate estimation of the projection parameters, but could be pared down further to provide initial and rougher forecasts. With time and resources, the estimates should be supplemented by other national and local analyses as proposed in Section 4.4.

Existing data could be specified and extracted now, while those referring to 2001 can be identified for quick subsequent extraction. In most cases, only ONS can provide the data, from commissioned tables.

General notes:

- Where Census is mentioned, data are for each subnational area, from commissioned tables if necessary. Where not specified, data are for both 1991 and 2001.
- Where samples are mentioned, data are for national areas only. In all other cases, data are for each District.
- In all cases, separate estimates are required for each ethnic group, subdivided for both the UK-born and the not UK-born.
- In all cases except fertility, separate estimates are required for males and females.
- For the base year, fertility and mortality, age refers to single years of age.
- Each analysis listed is indicative, not fully specified.

Base population

1. Census-based tabulations of residents, by age.
2. Students: District of residence by District of study, by age.
3. Armed forces and their dependants, by age.
4. An age-specific mapping of 1991 Census groups to 2001 Census categories, for 1991 Census data.

Fertility

1. Age-specific fertility, and propensity to change ethnic group for the newborn, from age of child, age of mother, and age of women (from the Census).
2. General fertility, derived from age of children and age of women.
3. Relationship between country of birth and ethnic group, from the Census. (To apply to birth registration counts, in order to estimate birth and death events and rates).
4. Each of 1–3 provided from each of Census, LS and LFS.

Mortality

1. Relationship between country of birth and ethnic group – see Fertility item 3.

Migration

1. Tabulations by five-year age groups and migrant status (flows between all Districts, within District, outside GB).
2. Estimate 1991–2001 national net overseas migration, and local net general migration, for each five-year age group, by measuring change in cohort size. Deduct 1991 population aged X from 2001 population aged X+5, adding expected deaths. Derive ten-year migration rates.

Economic activity

1. Tabulations by five-year age groups between 16–19 and 70+, economic activity (yes/no). From Census. From LFS each year during the 1990s.

Household formation

1. Household population, five-year age groups between 15–19 and 75+, household membership (as DETR). From Census, LFS, and LS.

2. Population not in households, age groups as in 1.

Administrative data

1. The development of name analysis procedures to be used with administrative data, and modelling their results to give estimates of populations by ethnic group for neighbourhoods.

2. The development and analysis of maternity records and immigration records as direct sources of demographic event information.

3. Calibration of non-Census data against the 2001 Census, including comparison of LFS and Census estimates of populations by ethnic group at mid-2001.

References

Aldous J, Bardsley M, Daniell R, Gair R, Jacobson B, Lowdell C, Morgan D, Storkey M and Taylor G (1999) *Refugee health in London: key issues for public health*, East London and the City Health Authority (The Health of Londoners Project): London.

Andelin J and Simpson L (2002) *POPGROUP: Population estimation and forecasting system reference manual*, City of Bradford Metropolitan District Council: Bradford.

Berthoud R (1998) Defining ethnic groups: origin or identity? *Patterns of Prejudice*, vol 32 no 2, pp 53–63.

Blackwell L (2000) Fragmented life courses: the changing profile of Britain's ethnic populations. *Population Trends* 101, pp 6–10.

Bradford Council (1996) *Labour force forecasts, 1996 round: technical report*, City of Bradford Metropoitan District Council Corporate Services, City Hall Research Section: Bradford.

Bradford Council (1997) *Household forecasts for Bradford District 1996: documentation*, City of Bradford Metropolitan District Council, Directorate of Corporate Services, Strategic Management Unit, Research Section: Bradford.

Bradford Council (1999) *Population forecasts for Rochdale 1999-2021: age, sex and ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford.

Bradford Council (2000) *Population forecasts for Bradford 1999-2021: age, sex and ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford.

Drew D and Simpson S (1981) *Bradford's black population: population forecasting and a fertility study*, Research Report MS/22, Department of Mathematics and Statistics, Sheffield City Polytechnic: Sheffield.

Harding S, Dews H and Simpson S L (1999) The potential to identify South Asians using a computerised algorithm to classify names. *Population Trends* 97, pp 46–9.

Haskey J (2000) Projections of the population by ethnic group: a sufficiently interesting or a definitely necessary exercise to undertake? *Population Trends* 102, pp 34–40.

Haskey J and Scott A (2001) More reliable minority ethnic population estimates and proportions for local areas: which areas to select and other choices to be made – an exploratory analysis. *Population Trends* 105, pp 16–27.

Luton Borough Council (2000) *Ethnic minority population in Luton*. Report to Planning Committee, 15 March 2000.

Manchester City Council (1999) *Planning Studies information note: estimating ethnic groups*, Manchester City Council, Planning Studies Group: Manchester.

Oldham Metropolitan Borough Council (1996) *Oldham's Social Survey 1995*, OMBC, Chief Executive's Department, Policy Unit: Oldham.

Oldham Metropolitan Borough Council (1997) *Population projections for Oldham*, OMBC, Chief Executive's Department, Policy Unit: Oldham.

Owen C (1993) Using the Labour Force Survey to estimate Britain's ethnic minority populations. *Population Trends* 72, pp 18–23.

Owen D (1999) *Making local estimates of minority ethnic group populations*. Paper presented in the Population Geography Research Group session on 'Forecasting Populations' at the Annual Conference of the Royal Geographical Society with the Institute of British Geographers, University of Leicester, 7 January 1999. (Available from Centre for Research in Ethnic Relations, University of Warwick, Coventry, CV4 7AL.)

Scott A, Pearce D and Goldblatt P (2001) The sizes and characteristics of the minority ethnic populations of Great Britain – latest estimates. *Population Trends* 105, pp 6–15.

Simpson S (1997) Demography and ethnicity: case studies from Bradford. *New Community*, vol 23 no 1, pp 89–107.

Simpson S (1998) Apportionment using counts of patients and electors, in Simpson S. (ed), *Making local population statistics: a guide for practitioners*, LARIA: Wokingham, pp 61–8. (An 'Estimating with Confidence' publication, available from LARIA, 2 Turnstone Close, Winnersh, Wokingham, RG41 5LQ.)

Simpson S and Close K (1996) *Ethnicity and household headship rates from the 1991 Census: availability of national and local data, CCSR Occasional Paper no 11*, University of Manchester CCSR: Manchester.

Storkey M (1999a) Refugee demographic data sources, in *Population Advice Note 98-5*, London Research Centre: London, pp 3–8.

Storkey M (1999b) Using the LS to investigate the changing fertility of different ethnic groups over time, in *Update – News from the LS User Group no 23*, Centre for Longitudinal Studies, Institute of Education: London, pp 3–4.

Storkey M (2000) *Methodology of the ethnic group projections 1999 Round*, London Research Centre: London.

Chapter 5

The starting population for population projections by ethnic group

Ludi Simpson

5.1 Introduction

Estimates of the future population are made by projecting forward from a date at which the population is known relatively accurately. At this base or starting date, the population must be known, or estimated, in sufficient detail to satisfy a projection model. Categories of ethnic group as well as age, sex and geography must be consistent with those for which fertility, mortality and migration or other measures of population change can be estimated.

This paper considers comparability between 1991 and 2001, which is important in order to:

- allow the estimated future path of population to be compared with a historical time series;
- estimate trends in the many demographic rates which can be derived from census data, to inform assumptions about future change;
- allow the success of pilot projections forward from 1991 to be measured against the outturn of the 2001 Census.

To summarise the paper's later conclusions on comparable relevant information from 1991 and 2001:

- Projections with a base year of 1991, and any comparison of 1991 and 2001 ethnic group statistics, should make use of the available estimates of complete 1991 Census data in order not to confound real change with a change in treatment of Census undercount which is expected to be estimated within the 2001 Census output.

- Direct comparison between 1991 and 2001 ethnic group categories will be approximate and in some cases misleading. A proportional conversion can be estimated and should be used where possible.

In considering a suitable base population for projecting forward populations by ethnic group, the following decisions must be made about the definition and derivation of ethnic group, geographical scale and time period. It is not the purpose of this paper to examine them and give recommendations, as this is done elsewhere, and is dependent on the purposes of population projections with an ethnic group dimension. The issues that should be resolved implicitly, or explicitly, by other papers are:

1) The detail of age, sex, economic activity, household structure, and other characteristics required;

2) The precise categorisation of ethnic group that will be used. From the 2001 Census, it will be possible to estimate any of the following:

a) Ethnic group as classified in Census output for England and Wales – currently expected to be sixteen categories: White: British, Irish, or Other White; Mixed: White and Black Caribbean, White and Black African, White and Asian, or Other Mixed; Asian or Asian British: Indian, Pakistani, Bangladeshi, or Other Asian; Black or Black British: Black Caribbean, Black African, or Other Black; Chinese or Other Ethnic Group: Chinese, or Other Ethnic group. These may be collapsed into broader categories in a variety of ways;

b) Polish, Cypriot, and other culturally cohesive origins, through use of Country of Birth categories;

c) Division of 'Indian' into Muslim, Hindu and Sikh to reflect better the different demographic experiences of the three groups;

d) Use of regional United Kingdom (UK) origins as further sub-divisions of ethnicity – English, Welsh, Scottish, Cornish.

3) Treatment of mixed ethnic group categories: a single 'melting pot', or specific combinations of non-mixed census categories, or allocation to a specific non-mixed ethnic group with likely closest affiliation. A single 'melting pot' does not make sense of the experience of the individuals with mixed parentage. The alternatives require estimation from the census data unless the census categories are accepted;

4) Subnational geographical scale: A geographically flexible estimation of smaller areas is implied by the Government's neighbourhood statistics strategy;

5) Geographical boundaries: frozen at the most recent census, or following boundary changes after the census;

6) Definition of the population: the usually resident population as used in ONS mid-year estimates is the feasible and expected choice. It is different from the legal concepts based on nationality and immigration status used elsewhere in Europe;

7) Special populations: Students or Armed Forces populations may be required separately, both for policy purposes and for demographic estimation of these two groups whose size and age-sex structure is often unusual, but stable over time;

8) Adjustments may be required to census data to provide a full mid-year population in the base year (as discussed later in this paper);

9) Base year: the above points assume national census year, but the use of survey and administrative data could provide a subsequent base year. However, it is likely that such estimates would incorporate the results of 'rolling forward' the populations from the last Census. Thus census year remains the relevant base year for practical purposes.

5.2 Comparing 1991 and 2001 ethnic group data: dealing with Census undercount

Procedures for the 2001 Census in the UK include a successful measurement of undercount and its characteristics, and a full allowance for this undercount within the census database before production of tabulations. The tabulations from the 2001 Census will therefore provide an estimate of the full population in census year, suitable as the starting population for projections (ONS, 2001).

In the 1991 Census output, however, undercount was only partially estimated. Table 1 indicates how Censuses (and surveys) tend to undercount minority groups more than the majority White group.

Race and ancestry have been asked in USA Censuses for many decades, and non-response has been consistently higher among the Black population and the more recently identified Hispanic population. In Britain, each resident's ethnic group was requested for the first time in 1991, with non-response also higher in the Black and South Asian (Indian subcontinent) minority groups, according to the Validation Survey report.

Country of birth has been a common proxy for ethnic group in studies in Britain, although it excludes the children of immigrants. In 1991, it allowed demographic comparison of census results with births in the years prior to the census, which confirmed non-response ten percentage points higher among children of those born in South Asia than children of all mothers born in the UK, and a still higher non-response for those born in African or Caribbean Commonwealth countries.

Race and ethnic group are social classifications that depend significantly on historical, political and legal conditions in each country, as evidenced from the many notes to the table above. Each group's non-response rate is likely to depend on local conditions including the ways in which census fieldwork adapts to cultural and social diversity. Nonetheless, high non-response among minorities in different countries may have common causes, including separation of communities from the nationally ruling culture of which the Census is part, and poor economic and urban housing conditions which are also associated with non-response (Simpson and Middleton, 1997; Simpson, 1999). Difficulties of attaining high response among minority race or ethnic groups are also evident in studies of sample surveys, and of the registration of voters (Smith, 1993).

Existing projections of the ethnic group composition of the population (see for example the paper 'Estimating methodologies and assumptions for projections' in this volume) have adjusted the 1991 Census to make it consistent with the ONS mid-1991 population estimates nationally and for each District. Some have applied the estimates of non-response in each ward, from the Estimating with Confidence project, to the Census age-sex structure of each ethnic group, to estimate the concentration of Black and Asian ethnic groups within the inner city areas which were most undercounted by the census (Simpson et al, 1997).

Most recently, those Estimating with Confidence estimates have been further refined to include an additional differential undercount between ethnic groups, applied within each ward-age-sex category. These are described below, and in more detail, by the current author (2002) and with Liz Middleton (1998).

The differentials embody the consistent evidence that minority and marginalised ethnic groups, the unemployed, and those renting private accommodation are most likely to be missed from the Census. For ethnic groups, the differential non-response adjustment – the additional proportion to be added to the census count – is

Table 1

Non-response rates. Ethnic group, race, country of birth

percentages

	GB 1991	GB 1991	USA 1990	USA 1990	GB 1991
	Census Validation Survey[a]	Demographic analysis, Country of birth of mothers of children aged 0-9[b]	Post-enumeration Survey[c]	Demographic analysis[c]	Labour Force Survey, Country of birth of head of household[d]
White	0.1	3.8[1]	0.7[2]	1.3[3]	14.1[1]
Black	5.2	21.6[4],17.5[5]	4.7	5.7	19.1[6]
Indian subcontinent	0.6[7]	13.3[8]			
Hispanic			5.0		
Other	1.4		2.4[9],12.2[10]		16.0

[1] *Born in the United Kingdom.*

[2] *Non-Hispanic White.*

[3] *Non-Black.*

[4] *African British Commonwealth countries.*

[5] *Caribbean British Commonwealth.*

[6] *All British Commonwealth.*

[7] *Pakistani, Indian, Bangladeshi.*

[8] *Pakistan, India, Bangladesh.*

[9] *Asian and Pacific Islander.*

[10] *Reservation Indian.*

***Sources:** [a] Heady et al, 1994. [b] Simpson, 1996. [c] de la Puente, 1993. [d] Foster and Bushnell, 1993.*

estimated to vary between White : Black groups : South Asian groups : Other groups in the ratio 1 : 6 : 4 : 3. In spite of the incomplete nature of the 1991 Census Validation Survey, it too showed this same regular pattern of higher non-response among ethnic minorities, and particularly the Black population.

Population estimates for each ward-age-sex category have been computed incorporating ethnic, employment and tenure differentials, so that the missing million – the undercount that biases the published 1991 Census counts – can now be put back into more applications. These estimates are available as 'SOCPOP' in the CD attached to Rees et al (2002), and to the academic sector via the MIMAS data service website (MIMAS, 2001).

SOCPOP provides data for ethnic, employment and tenure categories:

- Census day 1991 population: ethnic group (ten categories), employment status (four categories), or tenure (four categories), as well as five-year age group and sex;
- for the 9,636 LBS ward areas in England and Wales, and the 56 Districts in Scotland;
- allowances for the missing million population;
- no transfer of students to their term-time address, where students are enumerated in 2001. For comparability, further adjustment will be required.

Even when applying the response rate for each age, sex and locality *equally* to all social groups, there are quite large differences in their estimated level of non-response (black bars in Figure 1). This is purely because White, employed and owner residents are less likely to be in the age groups and places that ONS and the Estimating with Confidence project estimated as experiencing high non-response.

The plausible differentials between social groups' response rates used in SOCPOP make a considerable further difference (white bars in Figure 1). Groups other than White are estimated to number from 6 per cent to 17 per cent more than their census count, unemployed nearly 15 per cent more than their census count, and private renters about 10 per cent.

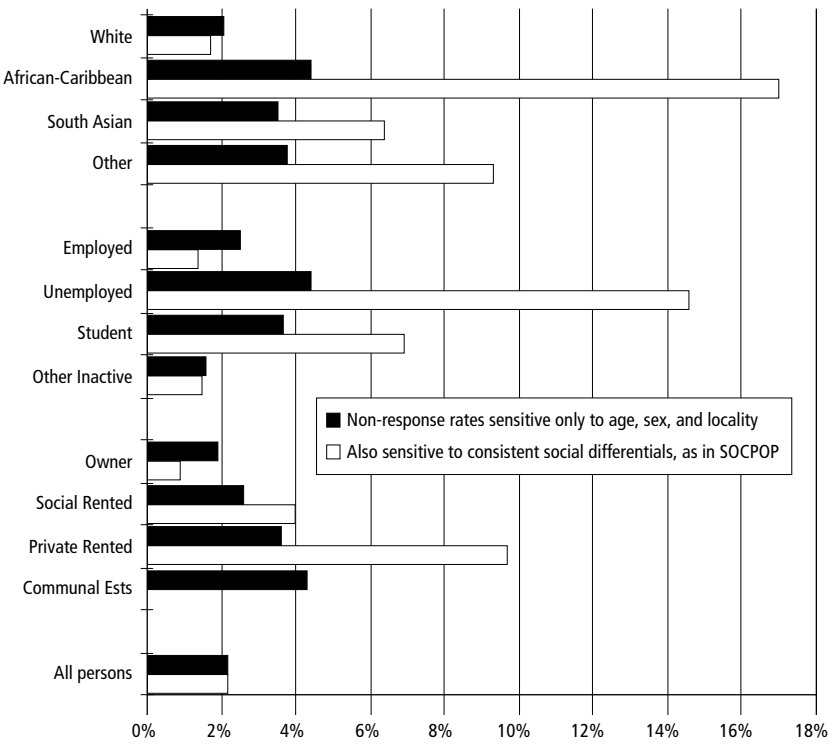
The differences make it important to use these full population estimates for analytical purposes; for example comparisons of unemployment rates between areas can be assessed for their robustness to census non-response.

The resulting population estimates are just that: estimates rather than certain counts. The differentials are plausible but not exact. In particular, due to lack of any relevant evidence, the differentials are applied to all areas in Britain equally and to all age groups equally. The 'South Asian' differential is applied equally to the Indian, Pakistani and Bangladeshi groups, though it is likely that their socio-economic and housing conditions would suggest a lower non-response for the Indian group.

SOCPOP data will be less biased than the published census results themselves when used as estimates of the 1991 national and subnational populations for each ethnic group, in the sense of more accurate on average, across many areas. But the error for any particular area is unknown and the possibility that it is significant should be borne in mind when considering changes in

Figure 1

England and Wales estimates of 1991 population from SOCPOP



population from 1991 to 2001. SOCPOP data for 1991 should be used, but key analyses should be duplicated using 1991 Census data as published in order to test the robustness of conclusions. For example, fertility rates and migration rates for specific ethnic groups can be derived from the Census. The rates, and in particular their variation between areas, may be affected by census undercount. To the degree that uncertainty about the level of undercount could explain the variation in vital rates, that variation should not be built in to a projection.

5.3 Comparing 1991 and 2001 ethnic group data: ethnic group classifications

This section looks only at the 10-category and 16-category classifications from the 1991 and 2001 Censuses respectively. There are arguments for using other classifications to enhance the categories for population projections, as mooted in the introduction. If pursued in practice, these other classifications would also require consideration of comparability between 1991 and 2001.

Table 2 shows a direct relationship between 1991 and 2001 Census categories, based on the likely affiliation of individuals when faced with both questions. The Table could be used to collapse 2001 Census categories to 1991 categories, and thus arrange 2001 Census data in a way that is suitable for comparison with 1991. Note that conversion in the other direction is not straightforward. The three 1991 'Other' categories are divided between more than one 2001 category, even when the five broad categories are used. Thus the 1991 category 'Black Other' is allocated partly to 'Mixed' and partly to 'Black or Black British', reflecting the introduction of a mixed category only in 2001.

From Table 2, a direct allocation between 1991 and 2001 categories could be made as in Table 3, retaining eight categories:

The assumption if using Tables 2 and 3 to compare across the censuses is that individuals do in fact respond to the different questions in the way indicated. It is an assumption that is unlikely to be true for all individuals. The problem would not be large for conversion of tabular data if mismatches for individuals in one direction tend to be 'cancelled out' by mismatches in the opposite direction, such that the error in the allocation of categories as indicated in Tables 2 and 3 are errors of approximation and become insignificant for relatively large populations.

Table 2

Conversion from 2001 Census to 1991 Census category

2001 Census broad category	2001 Census fine category	1991 Census category
White:		
	British	White
	Irish	White
	Other White	White
Mixed:		
	White and Black Caribbean	Black - Other
	White and Black African	Black - Other
	White and Asian	Other - Asian
	Other Mixed	Other - other
Asian or Asian British		
	Indian	Indian
	Pakistani	Pakistani
	Bangladeshi	Bangladeshi
	Other Asian	Other - Asian
Black or Black British		
	Black Caribbean	Black - Caribbean
	Black African	Black - African
	Other Black	Black - Other
Chinese or Other Ethnic Group		
	Chinese	Chinese
	Other Ethnic group.	Other - other

Table 3

Eight ethnic group categories compatible in both 1991 and 2001

Compatible category	1991 categories	2001 categories
White	White	White – all three subcategories
Black Caribbean	Black - Caribbean	Black or Black British – Black Caribbean
Black African	Black - African	Black or Black British – Black African
Indian	Indian	Asian or Asian British – Indian
Pakistani	Pakistani	Asian or Asian British – Pakistani
Bangladeshi	Bangladeshi	Asian or Asian British – Bangladeshi
Chinese	Chinese	Chinese or Other Ethnic Group – Chinese
Other	Black – Other	Mixed – all four subcategories
	Other – Asian	Black or Black British – Other Black
	Other - Other	Asian or Asian British – Other Asian
		Chinese or Other Ethnic Group – Other Ethnic Group

However, Tables 2 and 3 will not be even good approximations, to the extent that the different range of categories given in the two censuses, elicit not a direct transfer from one category to another, but a considered choice among several categories, the result of that choice reflecting factors that are not measured in the census. For example, the cultural and ideological affiliations of those people who declare mixed parentage in 2001, will have led many of them to identify with a specific ethnic group in 1991, rather than one of the three 'Other' categories as assumed in Tables 2 and 3.

This uncertainty suggests a probabilistic allocation between categories, by estimating the proportion of those self-identifying in a 1991 category who would choose each of the 2001 categories, and vice versa. Some information to estimate such proportions could come from the 1991 Sample of Anonymised Records, by comparing the ethnic group of parents and children. The update of the Longitudinal Survey (LS) will reveal the real choice of ethnic group categories made by the same individuals in 1991 and 2001. This comparison from the LS would include any changes of affiliation due to other factors, associated with age and period, which may not be insignificant as shown by other papers in this volume.

In conclusion, the data from the 1991 Census and 2001 Census may be compared using suitable allowance for census undercount and students, and matching the categories from each Census. Evaluations of projections from 1991 can be evaluated using data from the 2001 Census, but will be inexact because of the inexact match of the ethnic group categories in the two censuses. Equally, changes in vital rates measured using data from each census or other sources using the Census categories, need to be calculated to inform the assumptions about future rates to use in projections, but will again be inexact due to the inexact match of the ethnic group categories in the 1991 and 2001 Censuses.

References

De la Puente M (1993) Why are people missed or erroneously included by the census: a summary of findings from ethnographic coverage reports (including discussion), in *Proceedings of the 1993 Research Conference on Undercounted Ethnic Populations*, US Department of Commerce Bureau of the Census: Washington DC, pp 29–79.

Foster K and Bushnell D (1993) Non-response bias on government surveys in Great Britain, in *Fifth International Workshop on Household Survey Non-Response*, Statistics Finland: Helsinki.

Heady P, Smith S and Avery V (1994) 1991 *Census Validation Survey: coverage report*, OPCS series SS no 1334, HMSO: London.

MIMAS (2001) http://www.census.ac.uk/cdu/Datasets/1991_Census_datasets/Area_Stats/Adjusted_data/Undercount_adjusted_census_data/

Office for National Statistics (2001) *A guide to the One Number Census*, ONS: Titchfield.

Rees P, Martin D and Williamson P (eds) (2002) *The Census data system*, Wiley: London.

Robinson J G, Ahmed B, Das Gupta P and Woodrow K A (1993) Estimation of population coverage in the 1990 United States Census based on demographic analysis. *Journal of the American Statistical Association*, vol 88 no 423, pp 1061–71.

Simpson L (1999) Statistical exclusion and social exclusion: the impact of missing data. *Radical Statistics* 71, pp 45–60.

Simpson, L (2000) Away with 1991 Census Undercount! EWCPOP, SOCPOP and MIGPOP. *BURISA*, no 142, pp 10–14.

Simpson S (1996), Non-response to the 1991 Census: the effect on ethnic group enumeration, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census, volume 1: demographic chararacteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 63–79.

Simpson S (2002) Dealing with the Census undercount, in Rees P, Martin D and Williamson P (eds), *The Census data system*, Wiley: London, pp 171–183.

Simpson S, Cossey R and Diamond I (1997) 1991 Population estimates for areas smaller than districts. *Population Trends* 90, pp 31–39.

Simpson S. and Middleton E (1997) *Who is missed by a national census? A review of empirical results from Australia, Britain, Canada and the USA, CCSR Working Paper no 2*, CCSR: Manchester.

Simpson S and Middleton E (1998) Characteristics of census undercount; what do we know already? Adjusting census output for undercount, in Simpson S (ed), *A One Number Census: proceedings of a research workshop, CCSR Occasional Paper no 15*, CCSR: Manchester.

Smith S (1993) *Electoral registration in 1991*, OPCS series SS no 1301, HMSO: London.

This is a blank page

Chapter 6

Fertility

Deborah Sporton and Paul White

6.1 Introduction

With the stabilisation of many of the most significant minority ethnic groups of the United Kingdom, natural change has become the main driving force behind population change. Primary immigration and return migration, in comparison, now often play a role that has been much reduced since the early days of mobility, 30 or 40 years ago. Although the ageing of most minority ethnic groups is now occurring, age structures still tend towards being youthful in comparison with that of the White population. Warnes (1996) showed that, in the 1991 Census, average age differences between the White and minority ethnic populations was as high as 15.0 years for Black African women; 17.4 years for Pakistani women; and 19.3 years for Bangladeshi women. Age differences between the White and 'Black Other' populations were even greater. Although mortality certainly plays an important, and increasing, role in determining the growth rates of minority ethnic populations, fertility is now arguably the prime determinant.

Estimates of future fertility rates have generally been associated with a high degree of uncertainty, as changes in the social and economic climate can influence reproductive choices over a relatively short period of time. Fertility assumptions for the United Kingdom (UK) normally incorporate higher and lower variant total period fertility rates (TPFRs, calculated for a particular period, and constituting the number of children a woman would produce if her fertility experience through her life course were to mirror that of women of different ages within the period under consideration). The magnitude of deviation between these TPFR variants is often relatively small, yet the outcomes can produce significant differences in forecasted population size (Daykin, 1986). For some minority ethnic populations, this problem of estimating future fertility is compounded as higher overall TPFRs (for

example in the Pakistani population) are subject to multiple and dynamic influences constituted from both regions of origin and destination. Nevertheless, for the more well-established minority ethnic populations of the United Kingdom, it can be suggested that the likelihood of major demographic 'shocks' (such as a sudden migration surge) is now relatively low, and that many aspects of natural change may be susceptible to successful prediction and modelling.

A number of problems, however, must be overcome in any attempt at estimating future fertility by ethnic group. These include issues over ethnic identification, both of parents and of children; the likelihood of overall fertility dynamics that may differ between minority ethnic populations and the majority; what may be termed 'geographies of fertility' reflecting different fertility patterns in different areas; and (of particular importance) concerns with the availability of data on which to construct profiles of fertility by ethnic group for a point in time, and indications of rates of change. Issues of the fertility of minority ethnic or foreigner populations have been of some controversy in a number of countries. French public opinion has been assailed with a number of stories of high fertility predictions for the Maghrébin population, including the suggestion that in the future Marianne (the symbol of France) will wear a veil (Le Figaro, 25 October 1985). One aspect of Jürg Haider's political rhetoric in Austria has been concerned with what he sees as high levels of fertility amongst the Turkish immigrant population. In both cases, political discourse has been based on crude predictions, making no allowance for the dynamics of rapidly diminishing fertility amongst the 'problematised' groups. Accurate prediction must be based on a more informed conceptualisation of the situation, such that the findings are not decided in advance in ways that pander to racist discourses. One aspect of such detailed analysis requires the focus to be placed on all population groups and not simply on those who have traditionally been thought of as 'immigrants' or 'ethnic minorities'. Everyone has an ethnicity.

The following discussion is in two parts. In the first, we consider some of the more sociological and cultural (as well as practical) issues raised by ethnic fertility projections. In the second, we deal with data sources and methods of data handling to yield such projections.

6.2 Reflections on difficulties in projecting fertility by ethnic group

6.2.1 Definitions of ethnic groups

Haskey (2000) provides a consideration of issues concerned in defining ethnic group in the UK 1991 and 2001 censuses, and the Labour Force Survey (LFS). A number of other commentators have drawn attention to these definitional problems (eg Ballard, 1996, 1998; Berthoud, 1998). Rather than repeat many of the points raised in such commentaries, this paper reflects only on certain issues relating especially to fertility.

It is particularly important to bear in mind that a working definition of ethnicity might be accepted as 'both a way in which individuals define their personal identity and a type of social stratification that emerges when people form groups based on their real or perceived origins' (Hiebert, 2000, page 235). It is therefore a social construction influenced by social context, but operationalised at an individual level: this operationalisation occurs, for example, when a respondent is asked (as in the census and LFS) to provide a self-identification. Nevertheless, the social context (in this case operating through the design of official forms) limits the options available and therefore circumscribes the respondent's effective choices.

Given the socially constructed, rather than essentialist, nature of ethnic categories, it is clear that there will be a continuing need to regard them as dynamic, and to consider changes in the identification of groups to be analysed. This is needed in order to enhance and retain any social plausibility the categories used may have, to give them political legitimacy in relation to British society as a whole, and to give them demographic credibility as bases for projection. As will be argued later in this paper, moving away from the exclusive use of ethnic group (as recorded in the LFS and census) to a mixed approach to fertility calculations (including the use of mother's country of birth data as recorded at birth registration) could help in the production of more subtle categorisations of ethnicity in the future. Related approaches were adopted in the Fourth National Survey of Ethnic Minorities (Modood et al, 1997), where both self-ascribed identity and origins were sought from those in the survey.

Clearly there would be major problems with a dependence on mother's country of birth information alone, which would return the pragmatic understanding of ethnicity as a variable in British life to the position it held before the introduction of an ethnic group question in surveys in the 1980s. Country of birth is not a surrogate for ethnic group, and becomes less so as time goes on: significant proportions of most of the non-White minority ethnic populations in Britain today were born in Britain, and that is particularly so for those in the child-producing age groups. At present, birth registration does not involve a question on the ethnic group of either the newborn child or the mother, although it is conceivable that such a question could be asked in the future. Such questioning relating to the child might be seen as politically unacceptable, since it would be categorising the infant with an identity that might then be seen as unchangeable over the life course (see below). Since 1969, the mother's country of birth has, however, been asked at birth registration, and it is this key variable that provides a way into data estimation. The census or the LFS can be used to provide cross-tabulations of country of birth against stated ethnic group for women in the fertile age groups. Such a cross-tabulation can then be converted to a probability matrix that can be used to convert fertility by country of birth to fertility by ethnic group. This was one of the approaches used by Rees and Phillips (1996) to obtain estimates of Total Period Fertility Rates for a number of ethnic groups for 1991. Since such cross-tabulations (at least from the census) can be viably produced at local levels (such as districts or London boroughs) there is also the possibility of making local data sets sensitive to local variations in the ethnic group / country of birth equation. This is a potentially important issue since, as argued below, there may be a geography of ethnic fertility.

6.2.2 The ethnicity of newborns

If we accept that ethnicity is a social construction, then it can be argued that a newborn child does not have an ethnicity – it is not something that the child has as an innate in-born characteristic. The child's ethnic affiliation develops through its early life under the influence of a range of circumstances at a continuum of scales from parental and home environment variables, through neighbourhood characteristics, and education, to the wider social context in which personal identifications and affiliations are constantly negotiated and renegotiated through social interactions. External labelling processes imposed on an individual may also play a role in instituting a situation in which an individual sees her/him-self in a certain way (possibly in agreement with the external label, possibly as a reaction to it). Ethnic self-identification is therefore a complex issue. In practice, the social theoretical position just stated, although of some possible relevance for all newborn individuals, is of greatest significance for children born to parents of different ethnicities. Where a child is born to parents whose own ethnic affiliations are the same these issues, of ethnic 'emergence' are likely to be less variable. Thus, for example, any presumption of a child's ethnicity derived from a knowledge of the ethnicity of both parents is likely to be on relatively sure ground, whereas a presumption of ethnicity for a child with mixed parentage is less theoretically secure. Even where parents are both of the same ethnicity, however, there may be external pressure to label their children in a different way: there is evidence, from the 1991 Census for example, to suggest that a number of parents who identify themselves as Black Caribbeans label their children as Black British. In the 2001 Census results, this problem will be reduced through the use of Black British as the higher order category with Black Caribbean as a subgroup within it, but similar inter-generational labelling issues could occur in the future for other ethnic groups – such as the possible emergence of a BBC (British-born Chinese) identity (Pang, 2000).

Where the parents of a newborn are of different ethnic identifications, the problems for the calculation of ethnic-specific fertility rates, and for longer-term projections, increase considerably. As Aspinall (2000) has suggested, intermarriage rates between different ethnic groups are likely to increase in Britain in the near future, particularly since evidence from other countries (such as Sweden and the USA) consistently suggests that intermarriage is more common for 'second' and subsequent generations than for first generation immigrants (Cretser, 1999; Liang and Ito, 1999). The issue then arises of the ethnicity that is ascribed to the child in mixed ethnic unions.

Coleman (1985), using data from the 1981 LFS, carried out important work on intermarriage in Britain. More recently Smith et al (1996) modelled inter-ethnic marriages using the 1 per cent SAR from the 1991 census, looking at the extent to which there was symmetry in pairings (as many men from group x marrying women

from group y, as y group men marrying x group women). This supplements the important work by Berrington (1996) on the general characteristics of inter-ethnic unions shown up in 1991.

In terms of projections, the significant issue is the ethnic ascription that is then given to the children. Although Berrington (1996, page 204) indicated that '... investigation ... of the reported ethnicities of the parent(s) and children in mixed ethnicity families may prove very illuminating', it appears that no major studies have recently been published on this subject. Evidence from earlier periods is very significant. Coleman (1994) showed that around 20 per cent of minority ethnic 0–4 year olds in 1991 were of mixed ethnicity in terms of parentage. Using the LFS before the category 'mixed' was dropped, Shaw (1988) identified the fact that around 40 per cent of children born to mixed marriage couples where one of the parents was White were themselves labelled as ' White' (rather than 'mixed') by their parents. Shaw's paper dealt both with the probabilities of ethnic intermarriage and with the labelling of children according to the ethnic group of both parents. For example, the probability of a 'West Indian' man having a 'White' wife was 0.17. The probability of a White woman and a 'West Indian' man labelling their child as 'White' was 0.35. The probability of any 'West Indian' male marrying a 'White' woman and having a child labelled as 'White' was thus 0.17 x 0.35 = 0.06. A 6 per cent inter-generational shift from a 'West Indian' to a 'White' label is of some significance for projections. For other possible parental groups, the shift was much lower: for an 'Indian' woman the probability of marrying a 'White' man was low (0.03), as also was the probability of such a couple labelling a child as 'White' (0.28), yielding an inter-generational labelling shift of less than 1 per cent.

These calculations have been based on a knowledge of the ethnic group of the father of the newborn children. The availability of such knowledge is generally very limited. A question on the father's ethnic group is not asked at birth registration, and would not be seen as plausible in situations where the father is not acknowledged. Clearly, there is a need for the monitoring of trends in these ethnic labellings within mixed unions, since the outcomes are important in terms of projections. Detailed evidence elsewhere in the world, and particularly in the USA, indicates that the patterns of ethnicity ascription for children in mixed unions are complex. For mixed unions involving Whites with American Indians the predominant ascription is of a White label (Eschbach, 1995), but for children with one Asian parent the ascription choices appear to be relatively arbitrary (Xie and Goyette, 1997). Status issues over particular ethnic labels seem to be of some importance.

It is likely that a probabilistic approach to the ascription of a child's ethnicity (based on the ethnicities of the parents) might be plausible, both for those born to parents of different ethnicities and to those where both parents share the same ethnicity (where the likelihood of a similar ascription to the child is high). A probability matrix would be needed, probably derived from census responses, but this matrix would then have to be used together with information on the mother's ethnicity (rather than that of both parents), because of the non-availability of paternal ethnicity in birth data.

6.2.3 The geography of fertility

In the discussion earlier of the relationship between mother's country of birth and ethnicity, mention was made of the fact that the relationships within such a table may be specific to a particular geographical area. In an area with recent in-migration a larger proportion of a given ethnic cohort of women will have been born abroad than in an area in which migration took place further in the past. The major minority ethnic groups of Britain have moved from high fertility parts of the world to the low fertility social context of Europe. Fertility rates of new migrants have generally been those of the native-born British population.

It can be suggested that where minority ethnic populations have settled in patterns of ethnic concentration, the social mechanisms preserving traditional modes of attitude and behaviour may be fostered. Where minority ethnic populations have dispersed more widely, then the permeability to new behaviours may have been greater. If this is the case, then it is possible that any projection of fertility of a particular ethnic group might need to take account of possibly higher fertility in areas of concentration and traditionalism than elsewhere.

As a very crude indication, we can consider, for each ethnic group, the ratio of the number of children aged 0–4, to the total number of women in each London borough, using the 1991 Census.

There is clear evidence of higher ratios in boroughs with high concentrations of particular ethnic groups. For example, the ratio of Bangladeshi 0–4s to Bangladeshi women is highest – at 0.35 – in Tower Hamlets, the same borough with the greatest Bangladeshi concentration in the city. Marian Storkey (in preparation) has modelled fertility by ethnic group across London in 1991, estimating Total Period Fertility Rates, and has reported similar findings. Indian fertility within suburban Outer London is above average in the boroughs of: Barking; Brent; Ealing; Greenwich; Hillingdon; Redbridge; and Waltham Forest – very largely echoing the areas of Indian population concentrations. Storkey suggests a Bangladeshi TPFR of 4.75 in Tower Hamlets compared with an average for Greater London as a whole of 4.0. Within-group variations from borough to borough may be strong: for example from 0.92 to 2.21 amongst Black Caribbeans, or 0.55 to 2.49 for Chinese. Such analyses suggest that ethnic fertility projections need to be made sensitive to local circumstances. The most obvious way to do so is through the derivation of local matrices of women's ethnic group by their country of birth for use with the registration data.

6.2.4 Time series

As indicated in the introduction to this paper, the scarcity of time series data for projections is a considerable problem when considering fertility for ethnic groups. Although researchers such as Rees and Owen have backwards-projected the data from the ethnic group question in the 1991 Census to generate ethnically-defined populations for previous years, incorporating such procedures into retrospective fertility calculations would probably be highly suspect. Fertility has clearly been dynamic amongst Britain's minority ethnic populations in recent decades (see Table 2), suggesting that projections should be based on recent data that is updated continuously.

6.3 Data sources for fertility projection

The introduction of the 'ethnic group question' in the 1991 Census provided, for the first time, detailed statistical information for different generations within the ethnic minority population. The census, however, does not provide demographic information on births and family formation patterns, but can be used indirectly to calculate child-woman ratios (Table 1). By providing data on the numbers of women by birthplace, within the reproductive age-range, census data can also be combined with birth registration data to calculate age-specific and total period fertility rates for a census year. Vital registration data on births are available according to the age and birthplace of the mother, allowing the estimation of age-specific fertility for first-generation minority ethnic women. The absence of a question on the ethnic group of the mother at the registration of births hinders the estimation of the fertility of second and subsequent generations. This is a serious limitation of this data source, as according to the most recent estimates, just under half of the minority ethnic population (non-White) was born in the UK (Schuman, 1999). Due to differences in the timing of migration to the UK between groups, however, the proportion of minority ethnic women aged 15–49 born in the UK varied considerably by ethnic group in 1991. For example, of all 'Other Asian' and 'Bangladeshi' women in the age group, those born in the UK represented only 10.6 per cent, and 11.1 per cent, respectively compared to 79.9 per cent of the 'Black Other' group' (Storkey, in preparation).

To overcome the problem of estimating the fertility of UK-born minority ethnic women, there have been attempts to allocate birth data into ethnic groups using estimates based on the Labour Force Survey (LFS) and the 1991 Census. The LFS provides information on the distribution of women according to both place of birth and ethnic group (see Shaw, 1988; Owen, 1995), from which, proportions can be calculated to assign births (from registration data) to an ethnic group. Rees and Phillips (1996) similarly calculated conditional probabilities of belonging to an ethnic group based on country of birth and applied these to births to calculate fertility by ethnic group. Minority ethnic women born in the UK, (the second or subsequent generations) can also be identified – although the allocation of births to these women is problematically based on assumptions about their fertility which are discussed later. The LFS has also been used to estimate age-specific fertility by relating the date of birth of women to those of co-resident children (Berthoud, 2001). This estimation, however, relies on the assumption that the women were the mothers of the children, and that the date of birth information provided is accurate. Berthoud also had to aggregate data from 13 years of the LFS to obtain adequate sample sizes for his calculations, thus foregoing the possibility of considering temporal trends.

An alternative data source on births that includes ethnicity are Hospital Episode Statistics which have been collated since 1995–96 by the Department of Health and in certain Health Trusts for considerably longer. As the majority of births take place in hospital – for example, 98.6 per cent in 1992 (Coleman, 1996) – the data can be used to analyse generational differences in fertility according to ethnic group. In Bradford, for example, hospital maternity statistics, available locally since 1977, have formed the basis for population projections (Bradford Council, 2000). Combined with vital registration data, the Bradford data set includes for the child: the date of birth; sex; and ethnic group, and for the mother: date of birth; residence postcode; number of previous live births, and, since 1995, ethnic group of the mother (although only in one of the Health Trusts). Data on maternities from hospital records are however subject to a number of limitations. The historical record (pre 1995–96) is highly variable between and within Health Care Trusts, both in coverage, classification of births according to ethnic group and accuracy. The ethnic group classification in Bradford, for example, was different to that of the census until 1995. While this data can be sensitive to local concerns and reflect important local ethnic patterns, there are issues concerning the extent to which they can provide universally comparative data. Although hospital records are created for all births, the coverage for home births and those that took place in private hospitals is poor.

For the purposes of projection, cohort rather than period measures of fertility provide a more stable indication of changes in the pattern of childbearing through the reproductive lifecourse and are therefore preferable for developing fertility assumptions. Past national population projections (Cooper and Shaw, 1993) therefore used data on reproductive intentions from the General Household Survey (GHS) to help prepare fertility assumptions. The General Household Survey has been running since 1971, albeit not for every year, and providing information on fertility by ethnic group. The potential of the GHS for developing assumptions about ethnic minority fertility patterns is limited by small sample sizes precluding the calculation of robust fertility rates by ethnic group. There are also problems associated with the estimation of fertility from intentional data as reproductive intentions are rarely realised. There is clearly a need for a British equivalent to the French Enquête Famille, a large (a quarter of all women enumerated) demographic survey run in parallel with, and linked to, the French population census. The Enquete Famille overcomes many of the limitations of British data sources by providing detailed information on each parent and each child that can be matched to census records.

Analyses of minority ethnic fertility have been reliant, to date, on a combination of data sources to prepare estimates (with differing margins of error) and to develop assumptions. However,

Table 1

Data sources for analysing ethnic minority fertility

	Place of Birth	Ethnic Group	Fertility
Census (Total population although under-enumeration of ethnic minority groups)	Place of birth of mother and father -Place of birth of children	-Ethnic group of mother and father -Ethnic group of children	(indirectly) -Resident children aged 0-4 years -Date of birth of children
Longitudinal Study (1% linked sample)	-Place of birth of mother and father -Place of birth of children	-Ethnic group of mother and father -Ethnic group of children	-Date of birth -Sex of child -Age of mother at birth
G.H.S. (Sample Survey)	-Place of birth of Head of Household	-Ethnic group of Head of Household	-Reproductive intention -Children in household
Labour Force Survey (Sample Survey)	-Place of birth of individual (M/F) surveyed	-Ethnic group of individual (M/F) surveyed	-Date of birth of women surveyed -Date of birth of co-resident children
Hospital Episode Statistics - from 1995-6 (Department of Health). **Maternities data from Health Trusts**	-Place of birth of child	-Ethnic group of child -Ethnic group of mother recorded by some Health Trusts	-Date of birth -Sex of child
Birth Registrations	-Place of birth of child -Mother's country of birth		-Date of birth -Sex of child

from 2001, detailed statistics on minority ethnic fertility should be available for the first time from a single data source, the Longitudinal Study (LS). Since 1971, 1 per cent of the population have been sampled and their records linked through to the 1981, 1991 and the forthcoming 2001 Census. Birth data in the Longitudinal Study have also been linked since 1971 to their census records and through record linkage, members of the sample can be retrospectively categorised according to ethnic group. The LS is therefore the only national data source for calculating fertility by ethnic group. However, for 1971 and 1981, sample sizes were too small to calculate robust fertility measures. Sample sizes in 1991 were larger yet Table 2 indicates that for some groups there were still less than 250 women in the sample resulting in estimates that are below those calculated using other data sources. Given the growing maturity of the minority ethnic population, the 2001 Longitudinal Study should provide a sufficient sample from which to analyse fertility over the reproductive life course for both first and subsequent generations and to inform assumptions for projections.

6.4 The analysis of minority ethnic fertility: developing assumptions for projection

The analysis of minority ethnic fertility has been limited by data constraints and has been predicated, in the case of those born in the UK, on assumptions about reproductive behaviour based on current estimates (Table 2). Table 2 reveals that current estimates of the TPFR upon which such assumptions must be based are subject to marked variation for some groups according to the data source used – highlighting some of the data issues discussed above. Table 2 does, however, reveal variations between ethnic groups that must be incorporated into a population projection model. It is hoped that the 2001 LS will provide more robust calculations of fertility upon which fertility assumptions can be based.

More important than current levels of fertility to the development of assumptions about future change, are time series data that have been used in projections to model change over time (Table 3). The general trend has been for fertility to decline during the 1980s with variations between groups in the speed of this decline with, for example, the Caribbean-born women displaying

Table 2

Summary of Total Period Fertility Rate (TPFR) Estimates by Ethnicity, 1991

	Source: Rees and Phillips (1996) based on birth data converted into ethnic group births using the 1991 census	Source: Berthoud (2001) based on the LFS
White	1.85	1.8
Black Caribbean	Black Total 1.58	1.8
Indian	2.09	2.3
Pakistani	3.80	4.0
Bangladeshi	5.32	4.7
Other	1.64	2.0

Table 3

Total Period Fertility Rates (TPFRs) by country of birth of mother, 1981–91, England and Wales

Mother's Country of Birth	Year										
	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Bangladesh	8.0	8.0	8.3	7.7							
Pakistan and Bangladesh	6.5	6.3	6.1	5.7	5.6	5.6	5.2	4.9	4.7	4.7	4.6
Pakistan	6.1	5.9	5.5	5.3							
Rest of Africa	3.4	3.3	3.1	2.9	3.0	2.8	3.2	3.6	4.2	4.1	4.2
India	3.1	3.0	2.8	2.8	2.9	2.9	2.7	2.8	2.4	2.2	2.1
Far East	1.7	1.9	2.1	2.2	2.0	1.9	1.8	1.9	1.7	1.7	1.7
East Africa	2.1	2.1	2.0	2.1	2.1	2.0	2.0	2.0	1.9	1.8	1.8
Caribbean	2.0	2.0	1.8	1.8	1.8	1.8	1.9	1.8	1.6	1.6	1.5
UK	1.7	1.7	1.7	1.7	1.7	1.7	1.8	1.8	1.8	1.8	1.8

(Source: Reproduced with permission from Storkey (in preparation) based on birth data)

lower fertility than UK-born women by 1989. Between 1981 and 1991 in England and Wales, TPFRs for women born in the UK were stable at around 1.7 or 1.8: in contrast, the combined rate for women born in Pakistan or Bangladesh fell from 6.5 to 4.6, whilst the rate for women born in India fell from 3.1 to 2.1.

By converting place of birth data, as in Table 3, to ethnic group data, the historical rate of decline in fertility can be modelled for the purposes of projection into the future. However this is reliant on a number of assumptions.

- Assumptions about the level at which fertility will cease to decline, and about prospects for an increase in the future.

- As the minority ethnic population matures, assumptions need to be made about levels of future fertility arising from mixed ethnic unions. Shaw (1988) and Owen (1995) used the Labour Force Survey to assign births to women according to mother's place of birth and ethnic group. For those of mixed ethnic group, both Owen and Shaw assumed that the majority of births were to West Indian mothers, and therefore assumed that this mixed group would have similar fertility.

- Similar assumptions need to be made about the fertility of UK-born minority ethnic populations. For example, in their projections, Shaw (1988) and Owen (1995), *assumed* that age-specific fertility rates of minority ethnic women born in the UK (non-White) would lie midway between the rates for all UK born women and those from the relevant ethnic group suggesting a process of convergence. Minority ethnic women, who were born in the UK and identified themselves as White, were assumed to have the same fertility as all UK-born White women. Owen (1995, page 28) however recognised as more UK-born cohorts enter the reproductive age groups that '*future estimates of this kind based on broad assumptions (rather than hard evidence) about the fertility rates of UK-born ethnic minority population would become increasingly uncertain*'.

We have suggested in the previous section that this hard evidence could take the form of detailed information on cohort differences in age-specific fertility for minority ethnic women which will be available from the Longitudinal Study following the census of 2001 and which could be cross-checked with other data sources. However there remains the issue of the ethnic ascription of subsequent generations and the offspring of inter-ethnic marriages which must also be built into the assumptions of future fertility. Data available in the Fourth National Survey of Ethnic Minorities (Modood et al, 1997) provides an indication of perceived group membership by family origins. Individuals who would otherwise fall into a mixed race category could be allocated to ethnic groups with associated fertility assumptions by refining the probabilistic approach advocated by Rees (2000).

As the minority ethnic population of the UK matures, so the various future fertility scenarios become more complex as other influences, notably social class, education, and region of residence, become increasingly relevant and the differences between subgroups within broad ethnic classifications become more significant. Possible future scenarios for minority ethnic fertility are identified which may be applied to different minority ethnic groups and which have been based on previously estimated trends.

Scenario A: Convergence. This scenario is drawn from the 'assimilationist' literature that suggests that, over time, with successive generations, minority ethnic fertility will converge with that of the UK-born White population. From Owen's (1995) analysis of 1981-1991 trends in fertility rates, this scenario would appear to approximate the fertility of the West Indian population, which has now fallen below that of the White population. The fertility of Chinese and Indian women has also stabilised and is approaching convergence.

Scenario B: This scenario is drawn from the literature which suggests that for some minority ethnic groups, notably the Pakistani and Bangladeshi populations, fertility rates will stabilise at a higher level than those for the White UK-born population as aspects of cultural retention promote larger family sizes (eg religion, marriage patterns) (Coleman, 1994).

Scenario C: Minority Group Status Effect. This scenario is based on work, notably of Goldscheider and Uhlenburg (1969) in the US,

which suggests that upwardly mobile, or high status, minority ethnic groups will display lower fertility rates than the national average in order to maintain their economic status. This was observed for the Jewish population in the US and may apply to some of the established White minority ethnic groups (eg of European origin) and to the Indian and Chinese populations who are upwardly mobile (Green, 1997; Heath and McMahon, 1997). Clearly there is a need to examine the interrelations between ethnicity, fertility and socio-economic mobility in the formulation of fertility assumptions.

Scenario D: Divergence. Fertility estimates for White minority ethnic groups have typically closely approximated to those for the White UK-born population (Owen, 1995). We suggest that a possible future scenario might be an increase in fertility for certain members of this group. The introduction of White asylum-seekers from high fertility regions, for example, may result in the future upturn of White ethnic minority fertility especially given the conditions of their settlement in the UK. However, given the issues raised earlier about data sources it is unlikely that any considerations of differential 'White' fertility can be built into predictive modelling at this stage.

6.5 Conclusions

We have highlighted certain broad definitional issues and conceptual problems with the projection of fertility by ethnic group, and the data requirements necessary to construct profiles of fertility by ethnic group and model change over time. In conclusion we identify, from the discussions, the following requirements that would enhance the accuracy of fertility estimates and dynamics for a projection model.

- There is a need to convert country-of-birth data for mothers to ethnic group. It would be possible to use finer ethnic group categories than those available in the ten-group census classification. Such finer categories could be obtained by the cross-classification of place of birth and ethnic group.

- There is a need for more research on the ways in which children are labelled with an ethnic group by their parents, with a particular need for such information with regard to children of mixed ethnic parentage.

- We recommend ideally that the mother's ethnic group should be recorded universally at birth. Working with existing data, however, the most fruitful approach would appear to lie in the application of probability matrices whereby the likely labelling of newborns according to ethnic group is predicted in aggregate from known outcomes according to maternal birthplace and ethnic group, based on census or survey data. We suggest that information on perceived ethnic group membership by family origins as found in the Fourth National Survey of Ethnic Minorities may assist in this task (Modood, 1997).

- There is a need to work with data at local scales (for example administrative districts, or boroughs within Greater London) since spatial variation needs to be taken into account in two dimensions:

  a. country-of-birth to ethnic group conversions (above) may not be nationally stable;

  b. fertility within given ethnic groups appears to vary according to the strength of the spatial concentration of the groups considered.

- The measures of fertility adopted in projections need to be cohort based rather than relating to period rates. This would enable the dynamics of fertility to be captured more accurately.

- There is a need to work with continuous registration data wherever practicable. Where 'benchmark' data are used (as for example to provide denominators for fertility rate calculations), there needs to be regular updating – for example through the adjustment of census-based benchmarks through reference to survey material.

- It is possible to devise methods for projecting fertility by ethnic group within the UK, using existing data sources. The Longitudinal Study from 2001 is likely, for the first time, to provide a sufficient sample from which to analyse minority ethnic fertility by place of birth, ethnic group with other social and economic indicators over time. However, the LS information will not be available until possibly 2003.

Finally, it will be necessary for researchers to convince themselves at all times that the benefits of such research and projections outweigh the potential disadvantages. The distinguished French demographer Hervé Le Bras (1998) has argued against the study of differential demographic behaviour by ethnic group in France – both on the grounds of impracticality and on the grounds of undesirability since such research can play into the hands of the extreme right by apparently 'reifying' and legitimating ethnic categories. There is an obvious need for sensitivity in any discussion of future fertility according to ethnicity in the UK.

References

Aspinall P (2000) The challenges of measuring the ethno-cultural diversity of Britain in the new millennium. *Policy and Politics*, vol 28 no 1, pp 109–18.

Ballard R (1996) Negotiating race and ethnicity: exploring the implications of the 1991 Census. *Patterns of Prejudice*, vol 30 no 3, pp 3–33.

Ballard R (1998) Asking ethnic questions: some hows, whys and wherefores. *Patterns of Prejudice*, vol 32 no 2, pp 15–37.

Berrington A (1996) Marriage patterns and inter-ethnic unions, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census*, v*olume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 178–212.

Berthoud R (1998) Defining ethnic groups: origin or identity? *Patterns of Prejudice*, vol 32 no 2, pp 53–63.

Berthoud R (2001) Teenage births to ethnic minority women. *Population Trends* 104, pp 12–17.

Bradford Council (2000) *Population estimates and forecasts for Bradford District to 2021: age, sex, ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford.

Le Bras, H. (1998) *Le démon des origines: démographie et extrême droite*, Editions de l'Aube: La Tour d'Aigues.

Coleman D (1985) Ethnic intermarriage in Great Britain. *Population Trends* 40, pp 4–10.

Coleman D (1994) Trends in fertility and intermarriage among immigrant populations in Western Europe as measures of integration. *Journal of Biosocial Science*, vol 26 no 1, pp 107–36.

Coleman D and Salt J (1996) The ethnic question in the 1991 Census: a new landmark in British social statistics, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census*, v*olume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 1–32.

Cretser G A (1999) Cross-national marriage in Sweden: immigration and assimilation. *Journal of Comparative Family Studies*, vol 30 no 3, pp 363–80.

Daykin C (1986) Projecting the population of the United Kingdom. *Population Trends* 44, pp 28–33.

Eschbach K (1995) The enduring and vanishing American-Indian: American-Indian population growth and intermarriage in 1990. *Ethnic and Racial Studies*, vol 18 no 1, pp 89–108.

Goldscheider C and Uhlenberg P R (1969) Minority group status and fertility. *American Journal of Sociology*, vol 74, pp 361–72.

Green A (1997) Patterns of ethnic minority employment in the context of industrial and occupational growth and decline, in Karn V (ed), *Ethnicity in the 1991 Census*, v*olume 4: employment, education and housing among the ethnic minority populations of Britain*, ONS, TSO: London, pp 67–90.

Haskey J (2000) Projections of the population by ethnic group: a sufficiently interesting or a definitely necessary exercise to undertake? *Population Trends* 102, pp 34–40.

Heath A and McMahon D (1997) Education and occupational attainments: the impact of ethnic origins, in Karn V (ed), *Ethnicity in the 1991 Census*, v*olume 4: Employment, education and housing among the ethnic minority populations of Britain*, ONS, TSO: London, pp 91–113.

Hiebert D (2000) Ethnicity, in Johnston R J et al (eds), *The dictionary of human geography,* 4th ed, Blackwell: Oxford, pp 235–8.

Liang Z and Ito N (1999) Intermarriage of Asian Americans in the New York City region: contemporary patterns and future prospects. *International Migration Review*, vol 33, pp 876–900.

Modood T et al (1997) *Ethnic minorities in Britain: diversity and disadvantage*, Policy Studies Institute: London.

Owen D (1995) *Ethnic minorities in Great Britain: patterns of population change, 1981–91*, NEMDA 1991 Census Statistical Paper no 10, Centre for Research in Ethnic Relations, University of Warwick: Warwick.

Pang K (2000) *Towards British Chinese futures: a social geography of second generation young Chinese people living in London*. Unpublished PhD thesis, Queen Mary College, University of London.

Rees P and Phillips D (1996) Geographical spread: the national picture, in Ratcliffe P (ed), *Ethnicity in the 1991 Census*, v*olume 3: social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration*, ONS, HMSO: London, pp 23–111.

Schuman J (1999) The ethnic minority populations of Great Britain – latest estimates. *Population Trends* 96, pp 33–43.

Shaw C (1988) Latest estimates of ethnic minority populations. *Population Trends* 51, pp 5–8.

Smith P W F, McDonald J W, Forster J J and Berrington A M (1996) Monte Carlo exact methods used for analysing interethnic unions in Great Britain. *Journal of the Royal Statistical Society*, Series C: Applied Statistics, vol 45, pp 191–202.

Storkey M (in preparation) *Projecting the future populations of ten ethnic groups in London from 1991 to 2011*. Unpublished PhD thesis, University of Southampton.

Warnes T (1996) The age structure and ageing of the ethnic groups, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census, volume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 151–177.

Xie Y and Goyette K (1997) The racial identification of biracial children with one Asian parent: evidence from the 1990 Census. *Social Forces*, vol 76, pp 547–570.

Chapter 7

External migration and ethnic group

John Salt

7.1 Introduction

This paper focuses on information about international migration to and from the United Kingdom (UK), with particular reference to minority ethnic populations. Most of the detailed analysis is derived from the Census which allows analysis for one year only (1990-91) and which has already been presented in Chapter 5 of Volume 1 of the four OPCS/ONS ethnicity volumes (Salt, 1996).

The ethnic group question in the Census focused attention on the conceptual links between minority ethnic populations and immigration per se. There is no concept of 'immigration' in UK legislation. Despite this, much academic and media comment sees immigrants and settled minority ethnic populations as one and the same, so that 'immigration' is presented by most commentators largely as a movement for settlement purposes of visible minority ethnic groups. The resulting confusion tends to obscure the complex nature of immigration to the UK, which involves a wide range of migrant types and motivation: labour, family reunion, settlement, asylum seeking, students, and a plethora of temporary moves. In this chapter, the term 'immigrant', based mainly on the definitions used in the various official statistics, refers to those who have recently arrived (usually within the last year).

A distinction thus needs to be made between 'immigrants' and 'ethnic minorities' (Coleman and Salt, 1992). British legislation does not recognise the concept of 'immigrant' (or 'emigrant'), although conventionally the term refers to a person who has recently moved home for a defined period (normally anything from 3-12 months). An ethnic minority is a culturally defined group, defined on such bases as religion, language, social or family structure, or value systems. Its members may not have been recently (or ever) mobile internationally. Conceptually, therefore, immigrants and ethnic minorities are very different and certainly not coterminous groups. Hence,

study of the migration patterns of ethnic groups tells us more about the latter than about the national patterns and processes of international migration.

One of the few things that migrant and minority groups have in common is that they are generally poorly captured in the official statistics - which in turn originate from a diversity of sources. Hence, the migration of ethnic groups measured by the census must be seen as only part of the contemporary UK immigration story (albeit an important and visible one), and by no means its entirety. Nevertheless, it has been the immigration of ethnic minorities, particularly from New Commonwealth sources, that has largely driven UK immigration policy during the last half century.

The paper is in three main parts. The first provides a review of the main statistical sources on international migration for the UK. The second presents information on the main migration patterns of the ethnic minorities derived from the Census. The third reviews data from the International Passenger Survey (IPS), especially using migration exchanges with the New Commonwealth (and Pakistan) as an (admittedly questionable) surrogate for 'ethnic' migration.

7.2 International migration data for the United Kingdom

7.2.1 Introduction

An inventory of sources that may yield data on patterns of international migration relating to the UK is a combination of partiality and complexity. This is a reflection of the rather haphazard way in which the present system has evolved in response to the development of immigration control during the 20th century. Most of the data are based on administrative systems, related to control rather than migrant numbers and characteristics. With the exception of the International Passenger Survey, they record only immigration. Only the Census in 1991 and the Labour Force Survey (LFS) record information on minority ethnic populations.

7.2.2 Stocks

(a) Census

The Census is of limited use overall as a source of data on international migration. Although it asks about place of residence a year ago, there is no question on nationality, nor year in which foreign nationals settled in the UK, nor country from which they came.

Birthplace is recorded but, in the absence of year of entry, is of little help. Ethnic group was asked for the first time in 1991. However, increasing numbers of people belonging to the minority ethnic groups are British citizens, many born in the UK, so the question's use as a migration source is limited.

(b) Labour Force Survey

The first LFS in the UK was conducted in 1973; the survey was biennial until 1983, and has been annual since 1984. Since its inception, it has been used increasingly by government departments to obtain information useful for framing social and economic policy. During the 1980s, it became widely used as a source of information on the labour force generally, and on the characteristics of the minority ethnic populations in particular.

The LFS is a major source of both stock and flow data on international migration. Before 1992, about 80,000 sample addresses were surveyed, with a response rate of 80-90 per cent. The new quarterly LFS survey, from 1992 onwards, consists of five 'waves', each containing about 12,000 households. One consequence of the change is that the data before and from that date are not directly comparable.

The survey includes all UK and foreign citizens. The nationality question means that all foreigners are included, and the LFS provides the only data source on European Union (EU) nationals working in the UK. The application of grossing factors means that one sample interviewee is factored up to about 300 people in total. This constitutes a major problem when dealing with foreign nationals. Both flow and stock figures may be below this threshold for individual nationalities, particularly when any disaggregation into migrant characteristics is attempted. Data are available on nationality, age, sex, occupation, industry, region of destination and ethnic group. Some information on the labour market characteristics of minority ethnic populations is published, but not on immigration

flows. It is possible to relate ethnicity to international migration, though this has not yet been done here. The LFS also provides information on year of entry into the UK. Little analysis seems to have been done with this variable.

7.2.3 Flows

(a) International Passenger Survey

The IPS is a voluntary, continuous sample survey conducted by the Office for National Statistics (formerly by the Office of Population Censuses and Surveys) that covers the principal air and sea routes between the UK and overseas, but has excluded those between the UK and Ireland until recently. It is the only demographic data source giving both immigration and emigration statistics: thus it has considerable value.

Most of those surveyed are short-term travellers, but a sub-sample of 'migrants' is identified. A migrant *into* the UK is a person who has resided abroad for a year or more and, on entering, has declared the intention to stay in the UK for a year or more. A migrant *from* the UK is a person who has resided in the UK for a year or more and, on leaving, has declared the intention to reside abroad for a year or more. These definitions coincide with those of the United Nations.

Data are available on citizenship, country of origin, destination region, age, sex, and occupational status. Unfortunately, the sample size of 'migrants' is small; around 2,500 in all. Hence, detailed analyses of migrant characteristics in relation to particular variables, such as country of origin or region of destination, have limited value because the standard errors are too high. Also, its definition is *based* on intention to stay, and there is no guarantee that those recorded as migrants do actually come or go for the specified period.

In an attempt to get round the problem of 'switching' (i.e. people who arrive or leave with the intention of being in or out of the UK for less than a year, but subsequently stay for more than a year), the Home Office makes an annual adjustment to the total (around 70-90,000 per annum). Unfortunately, this adjustment cannot be made to the figures for migrant characteristics by origin, age etc. In addition, an estimate is made of the flow between the UK and the Irish Republic. Thus, the best estimate of total in, out, and net international migration, is derived by adding together these three elements: the IPS sample; visitor switchers; and the UK-Ireland flow. It is this sum that is published by ONS as its estimate of international migration and used in its calculations of population estimates and projections.

The IPS does not allow the identification of ethnic groups, although it is possible to make some inferences on the basis of aggregated countries of origin and destination.

(b) Labour Force Survey

The LFS provides transition data on immigrants to the UK, by asking for address one year ago. It does not provide flow data. Because of small sample sizes, breakdowns showing the characteristics of individual nationalities are rarely possible. Only for the major national groups (such as Irish), are estimated total numbers of immigrants available. An attempt was made by OPCS in 1994 to compare IPS inflow figures with the LFS data subset where individuals were living abroad one year ago for 1987-91. The results suggested that either the IPS overestimated, or the LFS underestimated, the number of persons with country of birth/citizenship in the Old and New Commonwealths.

(c) Home Office data on immigration control and settlement

Home Office statistics result primarily from enquiries made in the process of immigration control, in order to ascertain the status of arrivals with respect to the Immigration Rules, and to ensure that only those who are entitled to enter and settle are allowed to do so. By definition, British citizens, EU nationals, and others who have a right of abode in the UK are exempt from control and appear only in the gross totals of arrivals. All non-EU nationals entering the UK must fill in a landing card for the purpose of immigration control.

Data on foreign nationals accepted for settlement are a by-product of immigration control, and they may be regarded as a main measure of longer-term immigration. Variations from year to year may reflect legal changes, changes in the Immigration Rules, and administrative difficulties, as well as pressure of demand. EU nationals are normally admitted to the UK for an initial limited period of six months, and are not accepted for settlement on arrival. No data on

such admissions are collected by the Home Office. EU nationals can, however, apply for settlement on removal of time limit.

All non-EU nationals entering the UK must fill in a landing card for the purpose of immigration control. For those granted entry for more than six months, the following data are recorded on computer files: date of arrival; length of stay granted; nationality; admission category; date of birth; sex; number of children; date of departure. The data are regarded by the Home Office as generally reliable, although there is uncertainty about matching embarkation with arrivals.

Information for all settlement cases is available on date of settlement grant, nationality, category of settlement and sex. For non-EU nationals accepted after initial admission for a limited period (now over 80 per cent of all acceptances), this information is computerised and also includes date of birth. In addition, similar information is available for all those granted an extension to a limited stay in the UK, eg as a student.

The numbers of acceptances by nationality, sex and category are published in *Control of immigration: statistics* (Home Office, annual)

(d) Adjustments to ONS migration data

Work undertaken by OPCS in 1993 showed that projecting forward the 1981 Census figures, using data on births, deaths and unadjusted migration, resulted in underestimates of the overseas-born population in the UK in 1991. An exact analysis could not be undertaken because the degree of undercounting in the 1991 Census was known to be greater overall than that in the 1981 Census, and the precise amount was not known for the overseas-born population. A similar exercise carried out in the 1980s using 1971 Census data suggested that no significant adjustments to the migration data needed to be made. However, this type of comparison is fairly crude, as the Census data relate to overseas-born people who regard themselves as usually resident in the UK - with little or no guidance as to how 'usually resident' should be defined - and do not relate to 'migrants' as defined in the International Passenger Survey (IPS).

Although some 80 per cent of asylum seekers are not granted asylum or extended leave to remain, nearly all of them need to be recorded as migrants as they stay at least a year while their application is decided and their appeal against refusal is heard. Such persons will be recorded as outward migrants when they leave the country, because they have spent at least a year here, and therefore they need to be recorded as inward migrants at some stage.

It had always been assumed that these persons were covered by the IPS. However, when 'asylum seeker' was introduced in the 1994 survey as a 'reason for migrating', it was found that only around a thousand or so of the migrants (based on a small number of contacts) were seeking asylum. Subsequent inquiries made at the ports showed that immigration staff accompanying asylum seekers to Baggage Reclaim discourage any 'outsider' from talking to the applicant before the latter is initially interviewed. Once the asylum seeker has been interviewed, and given temporary admission to the country or detained while their application is considered fully, the person is unlikely to cross the IPS counting line again or might do so 'after hours'.

Therefore, for 1995 onwards, an adjustment has been made for port asylum seekers. Allowance is made for the fact that some are already in the IPS and some (currently an estimated 10 per cent) do not stay for a year or longer because they are refused and removed within a year, for example, because they had entered the UK from a safe third country.

An adjustment for these asylum seekers has been made for 1989 onwards to reflect the substantial increase in the number of such persons. Again, allowance is made for the fact that some (currently an estimated 10 per cent) do not stay for a year or longer, or were likely to have initially entered the country for a year or longer and therefore should not be included again.

An adjustment for these persons, mainly visitors, but including short-term students who switch, has been included for 1981 onwards to reflect the substantial increase in such persons. Most of the switching is due to marriage, or becoming a student. Information on non-European Economic Area nationals switching category after entry and being allowed to remain in the country for a year or longer is available from the Home Office database. However, the latter is not perfect, and in many cases the previous category is blank. This normally means that the person initially entered as a visitor or short-term student, whose arrival is not recorded on the computer. However, it can

also mean that the relevant information has not been entered.

As a result of a review in 1996 of the methodology used for deriving all the adjustments to the IPS migration data, including gauging the feasibility and magnitude of various types of switching, it was decided that the adjustments made previously for 'non-asylum switchers' were likely to be too large. Taken over a number of years, this overestimation offset the omission of adjustments for port asylum seekers for years prior to 1995. The current methodology is more likely to underestimate, rather than overestimate, the numbers of 'non-asylum switchers'. This is deliberate, and reflects the wish not to over-adjust the IPS data, and the fact that no adjustment can be made for those inward intending migrants who actually stay for less than a year.

The number of non-asylum switchers of USA or Old Commonwealth nationality is known, and the arbitrary assumption is made that there are equal numbers of British citizens visiting these countries who switch, i.e. stay for a year or longer. This is the only adjustment made to the outflow data.

It is estimated that such people led to a gain of around 50,000 in both 1991 and 1992, 38,000 in 1993, 45,000 in 1994, 48,000 in 1995, 40,000 in 1996, 32,000 in 1997, 46,000 in 1998, 73,000 in 1999 and 96,000 in 2000, net of persons leaving the UK for a short-term period who stayed overseas for longer than originally intended. This adjustment needs to be added to the IPS migration data to give a better estimate of *actual* net migration. Similar estimated adjustments ranging from around 10,000 in 1982 to 50,000 in 1991, and averaging around 23,600 per year, should be made to the IPS data for 1982-93.

(e) National Health Service Central Register

The NHSCR allocates a new NHS number and prepares a migration record for all new patients with a place of birth stated to be abroad. No information on the actual resident status of the person is available. The data may include some short-stay visitors and settled immigrants who have been in the UK for some years, but did not register with a doctor earlier and would be counted as recent immigrants upon registration. NHSCR data on migration from abroad include only age and sex. They are particularly incomplete as a record of emigration, since most people leaving do not inform their doctors. Further, the NHSCR does not use the same definition of international migration as the UN. These weaknesses mean that the NHSCR is little used as a source of migration data.

7.2.4 Naturalisation and citizenship

The UK has traditionally had a more relaxed attitude towards naturalisation than many of its neighbours, mainly due to past colonial influences. Most grants are discretionary rather than by entitlement. Data are available by former nationality, type and basis of grant, and are published in a *Statistical Bulletin* (Home Office, annual).

7.2.5 Evaluation

At first sight, the UK appears rich in international migration data. There is some scope for cross-checking between sources, though most of them should be regarded as complementary with one other. The IPS is a data source giving comparative immigration and emigration data, and conforming to the UN definition; it is also unique in Europe. When supplemented with visitor switcher and Irish flow data, it is probably an adequate basis for recording total numbers of migrants for the UK as a whole, together with breakdowns into broad categories. But the small sample size means it cannot be used for any detailed analysis. If the number of passengers continues to rise, special measures will be required to maintain the sample size of migrants. The Labour Force Survey also suffers from problems of small sample size, though not as acutely as the IPS. The LFS has unrealised potential as a source of data on stocks of foreign population and labour. The administrative nature of Home Office statistics means that no direct comparisons can be made with most other sources, although some comparisons are possible with IPS data.

7.3 Ethnic group and country of birth

7.3.1 Population

The 1991 Census recorded a non-White population of 3.02 million, 5.5 per cent of the total population of Great Britain (GB) (54.8 million); about 3.7 million were born outside the UK, of whom 2.14 million (57.2 per cent) were

White, and 1.6 million were non-White. While only 4.1 per cent of the White population were born outside the UK, this was the case for over one half (53 per cent) of the non-White population. Overall, 45 per cent of those born outside the UK were from the New Commonwealth, 34 per cent from Europe (including the Irish Republic).

There were considerable differences between the minority ethnic populations in the proportions born in/outside the UK. Around three-quarters of Asians and Chinese, and nearly two-thirds of Bangladeshis, were born outside the UK. In contrast, over one half of those from the Black Caribbean (53.7 per cent), and most (84.5 per cent) of the Black Other category, were born in the UK. Analysis of ethnic groups by country of birth shows the expected pattern of most Black Africans, Black Caribbeans, and those from the Indian subcontinent, who were not born in the UK, being born in the New Commonwealth. This was also the case for the Chinese and Asians, but with substantial proportions born in non-Commonwealth Asia (probably from the People's Republic of China via Hong Kong for many Chinese). The Other ethnic group, not surprisingly, shows the widest dispersion of birthplaces.

7.3.2 Immigration

Of the total number of immigrants who moved to live in Great Britain in the year before the 1991 census (326,398), 59 per cent were born outside the UK. Since nationality is not recorded in the census, there can be no indication of how many of these are returning Britons. However, the proportion is broadly similar to the 56 per cent of immigrants to the UK who were foreign nationals, recorded in the IPS.

Europe was the main source region (54,124), accounting for 28 per cent of all immigrants. Countries of the New Commonwealth provided 24 per cent (45,757) of those born outside the UK, those from the Old Commonwealth just under half that number. Overall, the Commonwealth as a whole was the birthplace of just over one-third of all immigrants.

7.4 Immigrants 1990-91 by ethnic group

According to the Census, a quarter (80,331) of immigrants were non-White. The largest groups was the Asians (18,160, 5.6 per cent of all immigrants), and Black Africans (15,536, 4.8 per cent). The comparatively small numbers of immigrants describing themselves as Black Caribbean reflects their long establishment in Britain, and indicates that the group is not involved in significant chain migration. The numbers of Black Other were also low, reflecting the UK birthplace of most of this group, also resulting in a small chain effect.

Some ethnic groups contained considerably above average (even for non-Whites) proportions of young people. This is particularly the case for Bangladeshis, Other, Black Other and Pakistanis. In contrast, Chinese, Indian and Black Caribbean immigrants contain relatively low proportions of children. In general, these differences reflect the recency of the immigration stream, but it is also a consequence of overseas events that have encouraged immigration towards Britain, including the growth in numbers of asylum seekers. At the opposite end of the age spectrum, several groups had relatively high proportions of elderly immigrants, Indians in particular, but also Pakistanis, Bangladeshis and Black Caribbeans. Black Africans and Asians, in contrast, had a low grey-haired immigrant representation.

In the main, the differences in age structure by ethnic group for the population as a whole were replicated for the two sexes, although there were some differences in detail. Immigrant Indian, Pakistani and Bangladeshi children were more likely to be boys than girls; this is also true of Black Caribbeans, but not the other two Black groups. Amongst the over 60s non-Whites, men were a smaller proportion than women, although the reverse was the case for Bangladeshis.

7.5 Migration trends in the 1980s and 1990s

Statistics from the IPS, for 1999, show an overall net gain of 181,500, a little higher than the year before and the largest net gain recorded. However, its make-up was different; one reason being the much larger adjustment made to take account of asylum seeker switchers. British citizens were no longer in balance, with a net loss of 8,400, but this was more than compensated for by a net gain of 189,900 or so non-British. However, compared with 1998, considerable changes occurred in the composition of the non-British. Flows of EU nationals were substantially down, as were those from the Old Commonwealth. These two groups

had been largely responsible for the steep rise in 1997-98. Numbers from the New Commonwealth rose in 1999, whereas they had fallen in 1998; the Other Foreign group continued to increase, to not far short of 100,000. Thus, the pattern of migration in 1999 was substantially different from 1998 even though the numerical outcome was similar. In the rest of this section, the figures quoted relate to unadjusted data, i.e. not taking account of category 'switchers'.

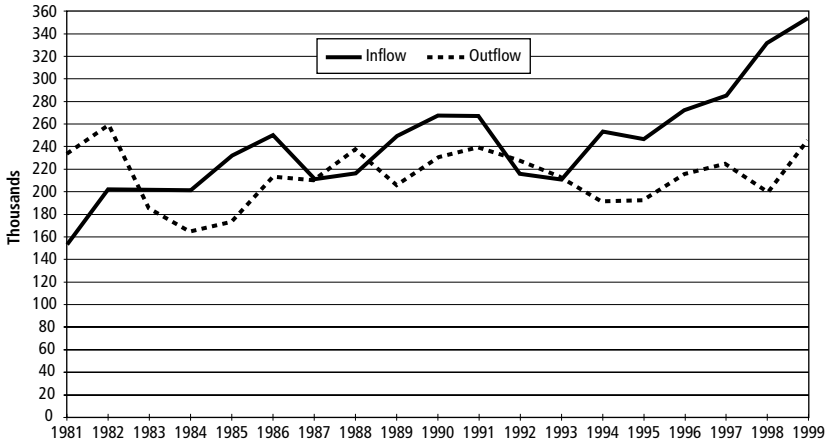
7.5.1 Recent Trends in Total flows

The pattern of overall inflow has fluctuated around a rising trend (Figure 1). Inflows of non-nationals (Figure 2) have consistently run ahead of those of the British, especially during the 1990s (Figure 3). The inflow of 354,000 in 1999 was the highest since records began in 1964, and was 22,000 higher than that in 1998. There was an increase of 5,000 in British citizen inflow and 18,000 in that of foreigners. The trend in outflows is less clear. The large increase in outflow during the 1979-83 recession, mainly by British nationals, was succeeded by comparatively low emigration levels in the middle 1980s. Until 1991, outflows increased steadily, among both nationals and non-nationals, although it was mainly the former who left. This was followed between 1991 and 1994 by a regular decrease in the outflow number of both nationals and non-nationals. This trend continued in 1995 for non-nationals, but not for the outflow of British nationals, which increased in number by about 10,000. In 1996, the total outflow increased, almost all of which was accounted for by more emigration among British nationals; however, the further rise in 1997 resulted from increased emigration by non-British people while the outflow of British citizens fell. In 1998, total outflow fell by 26,000, the bulk (20,000) of which was accounted for by a fall in emigration by British citizens. 1999 was different. Total outflow rose by 46,000 to 245,000, the highest figure recorded. The novelty was a sharp rise in the outflow of non-British, by 42,000 to 130,000, so that, for the first time, emigration of non-nationals exceeded that by the British (whose outflow rose slightly to 115,000).

Net flows were negative for the period 1975-83 (with the brief exception of 1979), then largely positive, except for 1988, 1992 and 1993 (Figure 4). However, from 1992 onwards, the net flow graph has been strongly upward, especially in 1998. The net gain of 60,000 in 1997 was the second highest in the period since 1979, that of

Figure 1

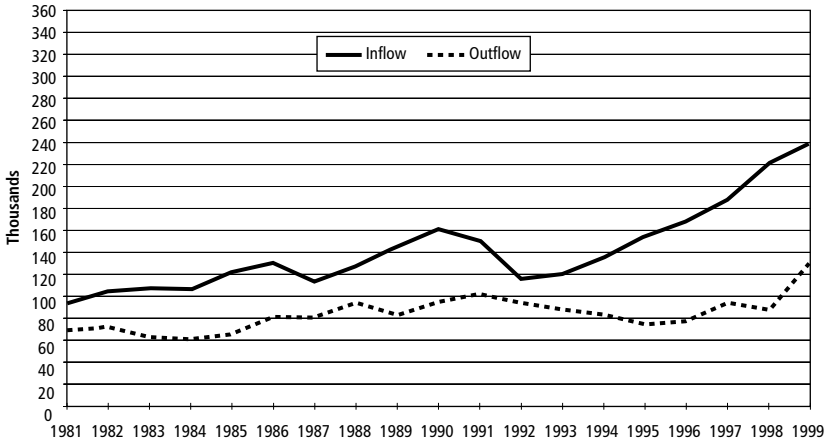
International Migration 1981-99. Total In- and Outflows



Source: International Passenger Survey, Office for National Statistics

Figure 2

International Migration 1981-99 Non-British Citizens In- and Outflows

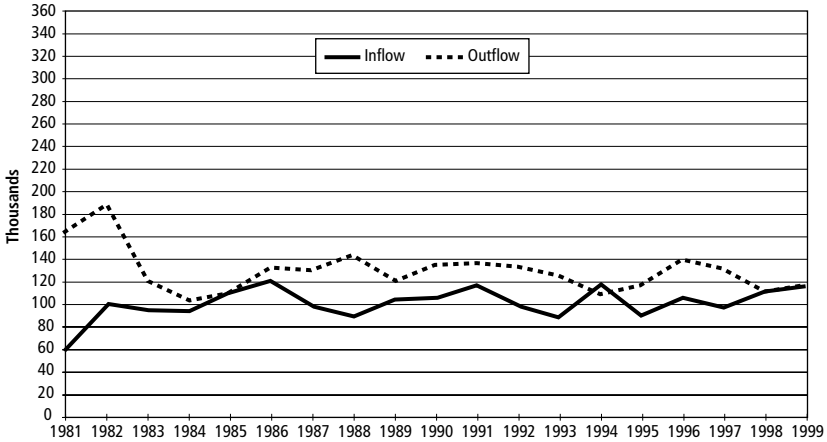


Source: International Passenger Survey, Office for National Statistics

133,000 in 1998 the highest ever recorded. In 1999, the net gain fell back to 109,000. The fluctuations depicted in Figure 4 for the most part reflect emigration by British people, among whom there was a net outflow for the whole period, apart from the break-even years of 1985, 1998 and 1999, and in 1994 when there was a net inflow of 10,000 British people.

Figure 3

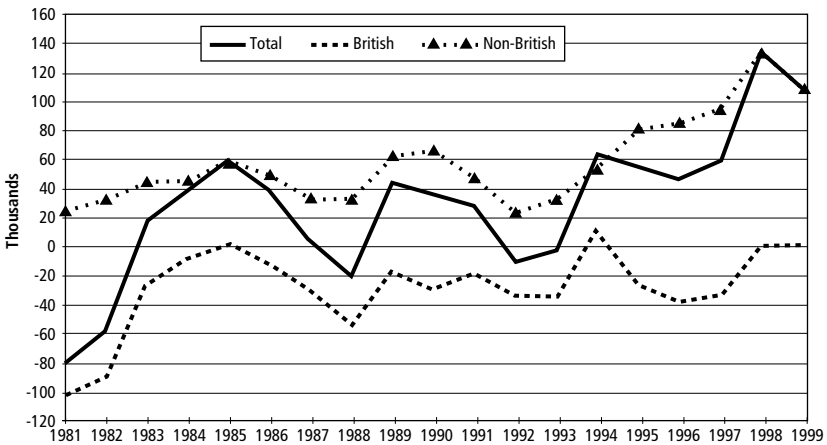
International Migration 1981-99
British Citizens In- and Outflows



Source: International Passenger Survey, Office for National Statistics

Figure 4

International Migration 1981-99
Net Flows by Citizenship



Source: International Passenger Survey, Office for National Statistics

Net flows of non-nationals have been positive throughout the period, the pattern broadly reflecting economic conditions, but with the amplitude of fluctuation less than that of British people. In 1998, the net flow of non-British rose by 41 per cent but fell by 19 per cent the following year. During the 21-year period 1979-99, net immigration of non-British citizens averaged 50,700 per annum, net emigration of British citizens 32,600, giving a net immigration average before adjustments of 18,100.

7.5.2 Flows involving people from the New Commonwealth and Pakistan

Table 1 summarises gross and net flows for the period 1981-99 for NCWP, Other non-British and British migrants. During the 19-year period as a whole, 4.619 million people came into the UK, 4.057 million left, and there was a net increase of 562,000. The equivalent figures for the NCWP countries are 698,000, 281,000 and 417,000 (Table 1), with annual averages of 37,000, 15,000 and 22,000. Put crudely, then, 'ethnic minority' net increases by migration have been in the order of 20,000 per year.

There has been no noticeable trend (Figure 5). Outflow has been within a very narrow band throughout, between 10,000 and 20,000 per annum. Inflow has fluctuated, mostly between 30,000 and 40,000, broadly in accord with the economic cycle, leaving net inflow varying from a low of 8,000 in 1988 to 31,000 in 1997.

7.6 The future

The future pattern of immigration by minority ethnic groups will depend on several major forces. New flows by foreigners will be heavily influenced by the evolution of government policy. At the moment, there are some signs that government attitudes may be shifting in the direction of greater liberalism, particularly in the case of labour migration. In the longer term, assumptions will need to be made on the basis of actual and projected trends in labour demand and supply. Currently, views on whether or not more immigrants will be required by European states to cope with demographic change are sharply polarised.

How far these changes will affect the minority ethnic population is difficult to say. Much will depend on how those populations now settled in Britain evolve. Patterns of family reunion are beginning to stabilise, at least among the older settled groups, but family formation that entails choosing marriage partners from the area of ethnic origin prolongs immigration chains.

We cannot assume that data on the immigration of particular ethnic groups is necessarily for settlement purposes. As former immigrant groups become settled, particularly from the second

Table 1

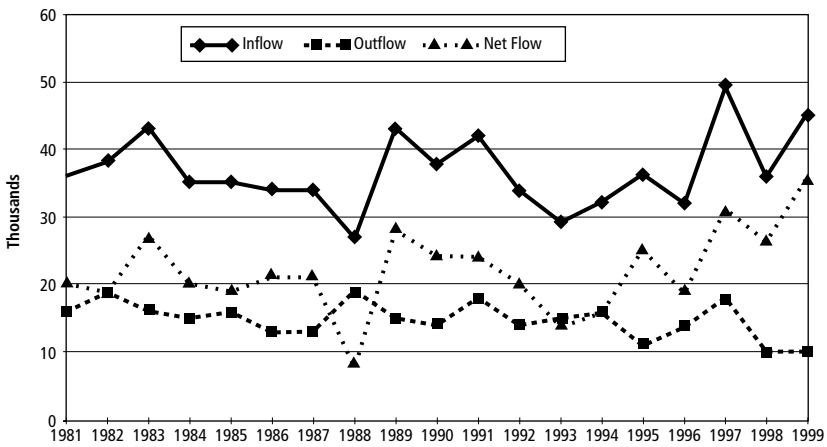
In-, Out-, and Net Flows to and from the UK, 1981-99

	New Commonwealth and Pakistan			Other Non-British			British		
	Inflow	Outflow	Net Flow	Inflow	Outflow	Net Flow	Inflow	Outflow	Net Flow
1981	36	16	20	57	53	4	60	164	-104
1982	38	19	19	66	53	13	98	187	-89
1983	43	16	27	64	47	17	95	122	-27
1984	35	15	20	71	46	25	95	103	-8
1985	35	16	19	87	49	38	110	109	1
1986	34	13	21	96	68	28	120	132	-12
1987	34	13	21	79	67	12	99	130	-31
1988	27	19	8	100	75	25	89	143	-54
189	43	15	28	102	68	34	104	122	-18
1990	38	14	24	123	81	42	106	135	-29
1991	42	18	24	108	84	24	117	137	-20
1992	34	14	20	82	80	2	100	133	-33
1993	29	15	14	91	73	18	89	124	-35
1994	32	16	16	103	67	36	118	108	10
1995	36	11	25	119	63	56	91	118	-27
1996	32	13	19	136	64	72	104	139	-35
1997	49	18	31	139	76	63	97	131	-34
1998	36	10	26	185	78	107	111	111	0
1999	45	10	35	194	121	73	116	115	1
Total	698	281	417	2002	1313	689	1919	2463	-544
Mean p.a.	37	15	22	105	69	36	101	130	-29

Source: International Passenger Survey

Figure 5

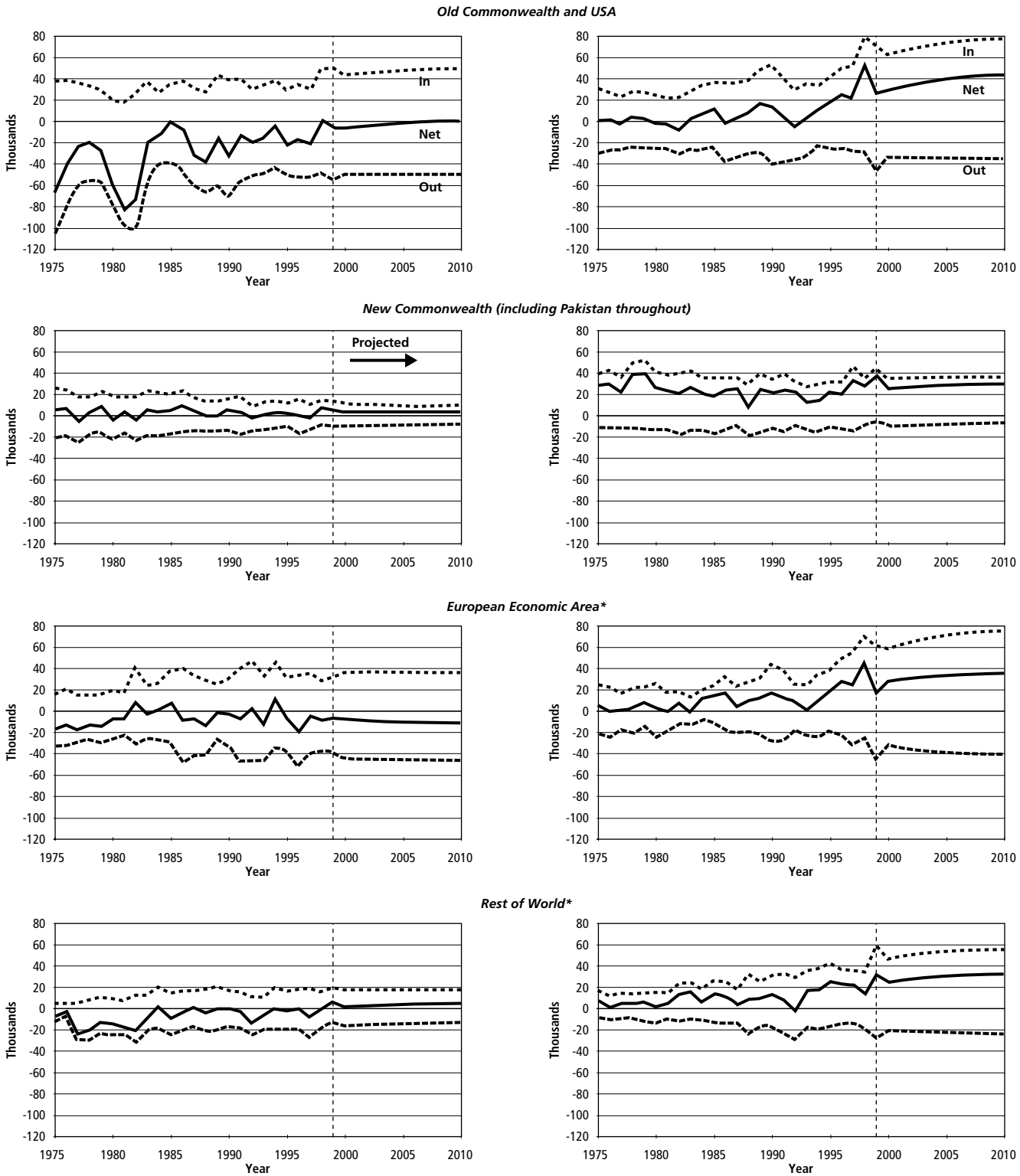
In-, Out- and Net Flows of New Commonwealth and Pakistan Nationals to and from the UK, 1981-99



Source: International Passenger Survey, Office for National Statistics

Figure 6

Figure 9.2 International migration to/from the United Kingdom by citizenship and country of origin/destination 1975-99 with projected trends to 2010



**Both categories exclude the Irish Republic*

Source: Government Actuary's Department

generation, many of their members will adopt a similar pattern of international migration to the host population. This will involve periods spent working abroad, from which return migration, as distinct from 'new' immigration is to be expected. This trend is likely to increase as fewer people generally opt for permanent emigration and temporary movements become the norm.

So where are we? In the light of the recent increases in net immigration, the latest (2000-based) projections from the Government Actuary's Department (Shaw, 2002) assume a net inflow for IPS migration of 100,000 compared with 65,000 from the previous (1998-based) projections. Allowing for visitor switching, the resulting overall assumption is for a net inflow into the United Kingdom of 135,000 persons per year, compared with the assumption in the previous projections of an annual net inflow of 95,000.

Figure 6 shows actual and projected net migration to/from the UK by citizenship and country of origin/destination to 2010. The scenario is for the continued net loss of British citizens to fall roughly to a balance and for a rising net gain for non-British citizens. Net migration for NCWP British citizens is projected to be just above zero; for NCWP non-British citizens at around 20,000 to30,000 per year - that is, at roughly the level of the last couple of decades.

References

Coleman D A and Salt J (1992) *The British Population*, Oxford University Press: Oxford.

Home Office (annual) *Control of immigration: statistics, United Kingdom*, TSO: London.

Home Office (annual) *Statistical Bulletin: persons granted British citizenship*, United Kingdom, Home Office: London.

Salt J (1996) Immigration and ethnic group, in Coleman D and Salt J (eds), *Ethnicity in the 1991 Census, volume 1: demographic characteristics of the ethnic minority populations*, OPCS, HMSO: London, pp 124-50.

Shaw C (2002) 2000-based national population projections for the United Kingdom and its constituent countries. *Population Trends* 107, pp 5-13.

This is a blank page

Chapter 8

Internal migration by ethnic group: data sources

Anne Scott and Lucy Vickers

8.1 Introduction

Measuring migration is not straightforward, as there is no compulsory system within the United Kingdom (UK) to record movements of the population, either into the country from abroad, or within the country. Therefore, in order to produce estimates and projections of the general population, at a subnational level, it is necessary to estimate migration using a combination of proxy and survey data. Measuring the internal migration of smaller subgroups of the population – such as by ethnic group – is more difficult again.

This paper examines the available data that can measure the migration of the minority ethnic populations within England and Wales. It assesses the existing data sources on internal migration and evaluates their quality. It also suggests the way that the data could be used to project internal migration by ethnic group. The paper is concerned solely with civilian migration.

8.2 Data sources

8.2.1 The Census

The Census of Population and Housing is the best source of data on internal migration within England and Wales. It includes a question that identifies individuals whose current address of usual residence differs from that of one year previously. All such people are classified as migrants.

Minority populations constitute only a small proportion of the population – 5.5 per cent of the total in England and Wales in 1991. Therefore, the Census is particularly important for identifying the patterns of migration of minority ethnic groups. Sample data may be adequate for estimating

migration flows between large areal units and for measuring the migration of large subgroups of the population. However, to estimate migration at a fine level of geographical or socio-economic detail, only the complete enumeration provided by the Census can provide sufficiently precise information.

The Census's strengths are:

- it covers all persons in the country and the measures produced are not affected by the imprecision that results from sampling;
- origins and destinations are recorded at the postcode level and can be aggregated to any geographical scale. (It is only because of confidentiality considerations that the most detailed data on migration flows cannot be released to users outside the Office for National Statistics (ONS));
- internal migration flows can be cross-tabulated with any of the many characteristics on which information is collected.

The Census's limitations are:

- the Census is conducted only once every ten years and therefore cannot be used for monitoring changing patterns of migration in intercensal years. (This may be especially significant if the Census is conducted in an atypical year in the economic cycle);
- the 1991 Census failed to enumerate 2.2 per cent of the total population and an even higher proportion of persons in the most mobile age groups;
- in addition, 6.1 per cent of the individuals who stated that they were living at a different address one year before the Census provided inadequate information on their former place of residence. These persons were therefore classified as 'migrants – origin not stated'.

In 2001, however, the combined impact of these two problems should be reduced since the 'One Number Census' methodology will mean that the undercount is estimated and included in the outputs.

The 1991 Census was the first to include a question on ethnic group. Publication of the results provided the opportunity for extensive analyses of the socio-demographic characteristics and geographical distribution of the ethnic minority populations. The numerous studies include several that have exploited the census data to investigate migration patterns. The most comprehensive of these is Tony Champion's, *Internal migration and ethnicity in Britain* (Champion, 1996). Other papers focus on migrants and the impact of migration on specific parts of the country (see, for example, Storkey, 2000a). A detailed discussion of the results of these studies is beyond the scope of this paper but some illustrative statistics on numbers of migrants by ethnic group are presented in Tables 1 and 2. Table 1 is a specially commissioned table provided by the London Research Centre (LRC). It is a matrix of flows of females of Chinese origin between the different boroughs of London and the rest of Great Britain. The figures shown are for all ages. Even at this aggregate level, the numbers of migrants in each cell of the Table are small. Moreover, it is clear from this Table that a similar matrix for inter-ward moves, or inter-ward moves broken down by age and sex, would consist largely of empty cells.

8.2.2 Health Service data

Internal migration of the total population is estimated using the proxy of patients moving and changing their doctor as they change address. The National Health Service Central Register (NHSCR) records the re-registration of patients as they move between former Family Health Service Authorities (FHSAs). These data are used to produce estimates of internal migration at the FHSA level (broadly county) for the population as a whole.

As a new initiative (see Scott and Kilbey, 1999 and Chappell et al, 2000), ONS has developed methods to utilise information from individual health authority patient registers to produce estimates of internal migration for all local and health authorities. To do this, ONS have obtained a complete copy of each patient register on an annual basis and by combining all 105 patient registers together, a register for the whole of England and Wales can be created. Comparing records in one year with those of the previous year – by linking on the NHS number – enables the identification of people who have changed their postcode. By comparing patient registers in two consecutive years, certain groups of moves that occur during the year will be missed. The data source cannot capture the movement of migrants who for one reason or another were not

Table 1

Migrant flows within London and between London boroughs and the rest of Great Britain, Chinese females, all ages, 1991 Census

		Borough of destination																																	
Origin		AA	AB	AC	AD	AE	AF	AG	AH	AJ	AK	AL	AM	AN	AP	AQ	AR	AS	AT	AU	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BJ	BK	Total
London borough																																			
AA	City of London	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
AB	Camden	2	66	3	2	9	6	9	1	2	3	1	5	8	17	0	20	3	6	0	3	3	2	1	5	0	0	3	1	1	1	1	0	3	187
AC	Hackney	0	3	21	1	6	7	0	0	3	5	0	4	1	10	1	1	0	0	0	2	1	3	3	1	0	0	0	0	3	3	1	0	1	81
AD	Hammersmith and Fulham	0	3	1	36	1	3	6	3	2	0	1	1	5	9	0	2	0	1	2	0	0	0	0	2	0	1	4	1	1	1	0	1	0	87
AE	Haringey	0	6	3	0	42	6	1	2	0	8	1	0	4	6	1	9	0	1	0	2	5	6	0	0	1	0	0	0	2	5	0	0	2	113
AF	Islington	0	2	4	0	9	40	0	1	2	0	3	2	1	5	0	5	0	3	1	1	1	5	2	0	1	0	1	0	0	2	0	0	2	93
AG	Kensington and Chelsea	0	4	0	13	4	2	30	2	4	0	1	2	4	18	0	3	0	6	2	4	1	0	1	3	0	1	3	1	2	0	2	0	2	115
AH	Lambeth	0	1	0	0	0	2	2	41	5	1	4	8	7	7	0	3	1	2	0	4	2	0	0	4	0	0	0	2	8	0	1	1	2	108
AJ	Lewisham	0	1	1	0	0	2	0	5	69	0	6	0	2	0	0	1	1	0	8	0	0	0	3	0	0	0	0	0	2	0	0	0	1	102
AK	Newham	0	1	9	0	2	2	1	0	1	51	2	2	1	6	3	2	0	0	1	0	2	1	2	0	0	2	0	0	1	6	3	1	1	103
AL	Southwark	0	5	2	1	0	3	2	5	12	0	28	2	3	3	0	3	1	2	1	0	1	0	0	1	2	0	0	2	0	1	3	0	0	83
AM	Tower Hamlets	0	1	2	2	3	0	1	4	7	6	2	52	2	4	4	2	0	0	0	0	2	0	0	0	0	1	0	0	1	1	0	0	2	99
AN	Wandsworth	0	4	0	3	0	0	0	3	2	0	5	1	53	5	0	2	0	0	0	2	1	3	0	0	0	0	2	5	8	2	0	5	0	106
AP	City of Westminster	0	20	3	7	5	3	17	3	3	1	6	6	3	72	0	18	0	6	0	2	10	0	0	0	2	2	0	0	3	1	2	0	0	195
AQ	Barking and Dagenham	0	0	0	0	1	0	0	0	1	0	0	3	0	0	10	0	4	1	0	0	0	0	0	0	3	0	0	0	0	4	0	0	0	27
AR	Barnet	1	4	0	1	5	2	2	0	0	2	3	2	1	13	0	97	0	4	0	0	4	1	0	5	0	0	2	0	2	1	0	3	0	155
AS	Bexley	0	0	0	0	0	0	0	0	4	0	1	1	0	0	0	0	31	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	43
AT	Brent	0	16	1	0	1	1	1	1	0	1	1	1	4	9	0	7	0	31	0	0	13	1	0	6	1	2	0	0	0	0	0	5	3	106
AU	Bromley	0	1	0	0	0	0	0	0	1	0	2	1	1	1	0	2	2	0	36	0	0	0	0	0	0	0	1	0	0	0	0	0	0	48
AW	Croydon	1	0	0	0	0	0	0	0	2	0	1	2	1	1	1	0	0	1	0	34	0	0	0	0	0	0	0	0	0	0	0	2	0	46
AX	Ealing	0	3	0	6	1	0	1	2	0	0	1	0	1	4	0	5	0	1	2	1	54	1	0	8	0	4	9	1	0	4	0	0	0	109
AY	Enfield	0	0	0	0	3	1	0	0	0	0	0	0	0	2	0	3	0	0	0	0	0	23	0	1	1	0	0	0	0	0	0	0	0	34
AZ	Greenwich	0	2	0	1	0	2	1	1	5	0	3	1	1	1	0	1	10	0	2	1	1	0	40	0	1	0	0	0	0	1	1	0	0	76
BA	Harrow	0	0	1	1	0	0	1	0	0	0	0	0	0	1	1	2	0	10	0	3	3	0	1	24	0	7	0	0	0	0	0	0	0	55
BB	Havering	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	17	0	1	0	0	0	0	0	0	21
BC	Hillingdon	0	2	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	3	0	0	1	0	31	0	0	0	0	0	0	1	41
BD	Hounslow	0	2	0	1	0	0	1	0	0	1	0	0	1	2	0	1	0	2	0	0	4	0	0	1	0	1	21	3	0	0	2	3	0	46
BE	Kingston upon Thames	0	2	0	0	0	0	1	1	1	0	0	0	2	0	0	1	0	0	1	2	1	0	0	0	0	0	0	29	1	1	1	1	0	45
BF	Merton	0	3	0	0	1	1	1	1	2	0	1	0	6	1	0	0	0	1	0	1	0	1	0	0	0	0	2	4	23	0	0	2	0	51
BG	Redbridge	0	0	1	0	0	0	0	0	0	7	0	2	0	1	5	6	0	0	1	0	1	2	0	0	1	0	0	3	0	40	0	0	2	72
BH	Richmond upon Thames	0	0	0	1	0	0	1	1	0	0	0	0	1	0	0	2	0	0	1	1	1	0	0	0	2	0	3	1	2	0	19	0	0	36
BJ	Sutton	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	5	0	0	1	1	0	21
BK	Waltham Forest	0	1	2	0	1	1	1	0	0	4	0	0	1	3	0	1	1	0	0	1	1	1	0	0	0	0	1	0	0	8	0	0	34	62
Origin																																			
	Same borough	0	66	21	36	42	40	30	41	69	51	28	52	53	72	10	97	31	31	36	34	54	23	40	24	17	31	21	29	23	40	19	11	34	1206
	Different borough	4	87	33	41	52	45	50	36	60	39	45	46	62	129	16	105	23	47	22	31	63	28	19	39	15	21	32	29	37	42	18	24	22	1362
	All London	4	153	54	77	94	85	80	77	129	90	73	98	115	201	26	202	54	78	58	65	117	51	59	63	32	52	53	58	60	82	37	35	56	2568
	(Other) Inner London	2	51	28	30	39	36	39	29	43	24	32	33	41	90	9	71	6	27	15	20	30	20	12	16	6	7	13	12	32	23	13	8	16	873
	(Other) Outer London	2	36	5	11	13	9	11	7	17	15	13	13	21	39	7	34	17	20	7	11	33	8	7	23	9	14	19	17	5	19	5	16	6	489
	Rest of GB	1	45	13	23	33	14	25	11	5	9	12	6	18	40	3	33	11	16	12	12	18	18	5	5	8	9	10	8	11	5	3	8	10	460
Destination																																			
	Different borough	2	121	60	51	71	53	85	67	33	52	55	47	53	123	17	58	12	75	12	12	55	11	36	31	4	10	25	16	28	32	17	10	28	1362
	Rest of GB	3	29	14	14	22	22	17	26	18	16	15	16	12	14	8	22	3	16	10	10	19	11	13	15	5	17	15	8	5	15	15	11	15	471

Source: 1991 Census, LRC Commissioned Table LRCT49. Figures exclude origin not stated and have not been adjusted for the Census undercount.

Table 2

Estimated migrant flows between district types for ethnic minority groups, 1991 Census

	Destination										
Origin	Inner London	Outer London	Principal cities	Other metro.	Large non-metro.	Small non-metro.	Industrial	Districts with new towns	Resort and retirement	Mixed urban/rural	Remote, largely rural
Black											
Inner London	23191	5748	207	110	129	100	201	135	128	374	115
Outer London	3734	13024	120	95	83	67	175	122	85	394	75
Principal cities	214	190	7593	701	112	49	87	44	22	139	49
Other metropolitan	153	103	657	4270	76	29	80	41	35	116	53
Large non-metro.	151	143	108	70	3620	56	155	21	38	211	66
Small non-metro.	104	94	75	39	48	1669	43	35	52	190	33
Industrial	130	164	77	69	143	33	2421	47	28	190	127
Districts with new towns	69	91	43	27	31	22	53	1174	27	174	30
Resort and retirement	86	57	21	22	16	41	24	17	688	97	53
Mixed urban/rural	285	316	102	122	209	166	199	138	73	3737	240
Remote, largely rural	79	65	41	31	57	51	84	52	47	272	988
Indian, Pakistani and Bangladeshi											
Inner London	10129	3176	157	243	157	83	194	90	123	369	86
Outer London	1721	16443	241	370	243	103	481	224	120	1151	115
Principal cities	120	383	11283	1873	220	123	269	179	44	291	85
Other metropolitan	199	468	1311	16875	329	137	573	233	57	423	143
Large non-metro.	130	361	227	303	6447	66	392	97	49	772	88
Small non-metro.	98	171	96	121	53	2172	113	83	55	220	55
Industrial	119	382	261	422	294	98	6394	141	55	472	102
Districts with new towns	57	158	102	111	66	43	178	1785	43	157	37
Resort and retirement	65	94	18	65	30	40	36	30	635	82	36
Mixed urban/rural	230	673	248	354	395	162	430	230	109	4595	142
Remote, largely rural	26	83	45	108	57	61	60	25	49	117	513
Chinese and other											
Inner London	9890	3283	136	123	127	145	168	105	143	536	103
Outer London	2058	11364	134	132	182	157	214	143	159	955	146
Principal cities	211	218	6510	837	132	96	173	129	65	226	86
Other metropolitan	99	134	694	3999	95	55	175	103	48	252	107
Large non-metro.	216	172	169	137	3931	95	299	56	59	424	148
Small non-metro.	205	170	117	57	112	2134	118	50	147	331	128
Industrial	100	208	119	149	241	66	2309	88	44	290	148
Districts with new towns	87	118	83	93	62	37	98	1356	39	256	58
Resort and retirement	125	146	49	72	52	135	51	34	1498	193	96
Mixed urban/rural	440	673	176	206	372	266	308	221	190	5326	310
Remote, largely rural	94	100	79	81	157	102	105	57	101	345	1697

Source: Rees and Duke-Williams (1997), pp 333-334

on the register in one of the two years, but who moved during the year. The largest group of these is migrants aged under one, but there are other groups. The NHSCR captures all these within year moves and so the more complete information from the NHSCR has been combined with the more geographically detailed information from the patient registers, to produce migration estimates for the total population for all local and health authority areas.

At present, the NHSCR and most health authority patient registers, lack data on ethnic group. It is possible that this situation will change. For example, provision has been made for the National Strategic Tracing Service (NSTS), a Department of Health initiative to hold all the details on individual health authority registers in a central facility, to hold ethnic group, should a reliable source be identified. ONS plan to receive a detailed migration extract from the NSTS, and

should ethnic group data exist, steps could be made to capture them in the future.

8.2.3 The ONS Longitudinal Study

The Longitudinal Study (LS) holds information on ethnic group of persons who were enumerated in the 1991 Census but does not collect details of moves or place of residence in intercensal years, i.e. it cannot be used to produce annual measures of internal migration of the population. Also, the LS does not include those who enter the UK as international migrants during the intercensal years.

8.2.4 Survey data

Most social surveys such as the General Household Survey and the British Household Panel Study interview only a small sample of the population. Since ethnic minority migrants represent a very small proportion of the total population, the migration estimates yielded by these surveys for the groups of interest are unreliable.

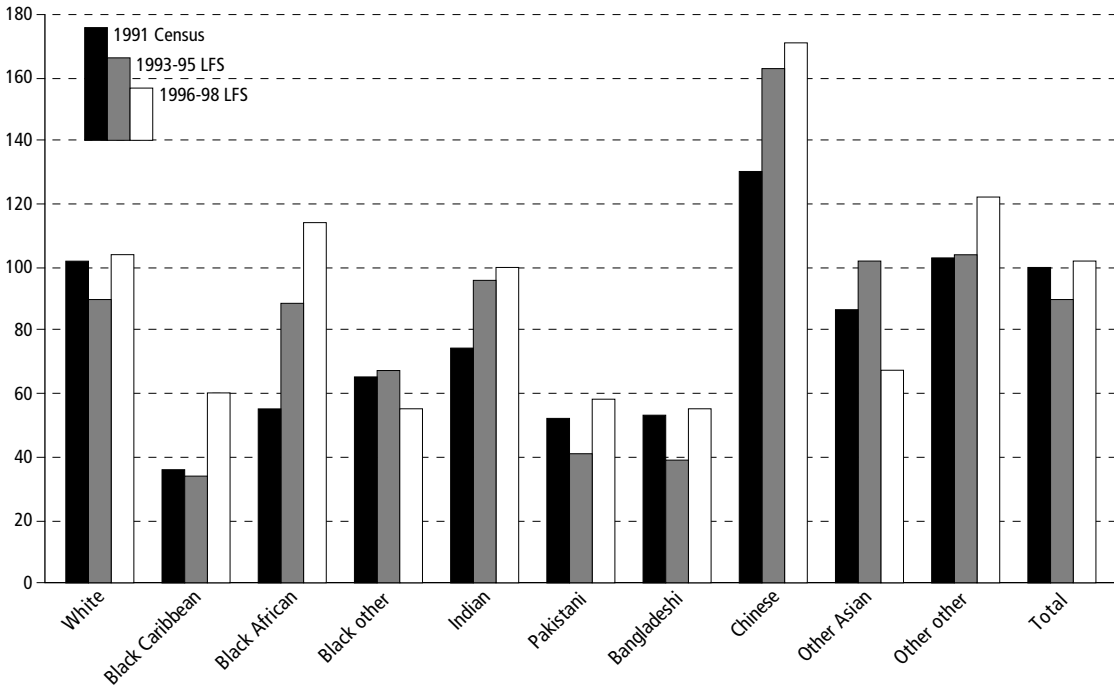
The only source that asks questions on both ethnic group and migration and has a fairly large sample is the Labour Force Survey (LFS). This survey currently interviews over 125,000 people (0.3 per cent of the population) in England and Wales each year. It therefore provides the best current source of information on the patterns of migration of minority ethnic populations since 1991. Figures 1 and 2 show some broad comparisons of migration patterns recorded by the 1991 Census and the LFS. The migration measure used is the standardised migration ratio (calculated in the same way as the standardised mortality ratio). The standard is age-specific migration rates in the 1991 total population. The LFS data are aggregated for two time periods, 1993–95 and 1996–98.

For most ethnic groups, the pattern of inter-regional migration revealed by the LFS data is consistent with the results of the full enumeration of the population in 1991. As in 1990–91, in the mid- and late-1990s, the Chinese population and people in the Other Other ethnic group were more likely to move between regions than the White population (Figure 1). Conversely, Black

Figure 1

Comparison of 1991 Census and Labour Force Survey migration data: migrants between regions

Standardised migration ratio



Caribbeans, Black Other, Pakistanis, and Bangladeshis were far less likely than the White group to undertake long distance moves. The LFS standardised migration ratios for inter-regional migration among the Black African and Indian populations suggest that the mobility of these groups has increased since 1991. However, the numbers on which these ratios are based are very small, hence the uncertainty surrounding their estimates is very large. For example, for Black Africans the ratios are based on 21 migrants in 1993–95 and 36 migrants in 1996–98. For Indians an average of 75 migrants was recorded in each three-year period. The LFS results for migration within regions also show similar relative levels of mobility as the 1991 statistics (Figure 2). The LFS ratios for the Black African, Black Other, Indian, Pakistani and Other Other groups indicate that their propensity to migrate over short distances has increased during the 1990s.

Migration estimates derived from the LFS, for small subgroups of the population, such as the minority ethnic populations, give rise to large sampling errors. The results do not therefore constitute reliable measures of migration trends.

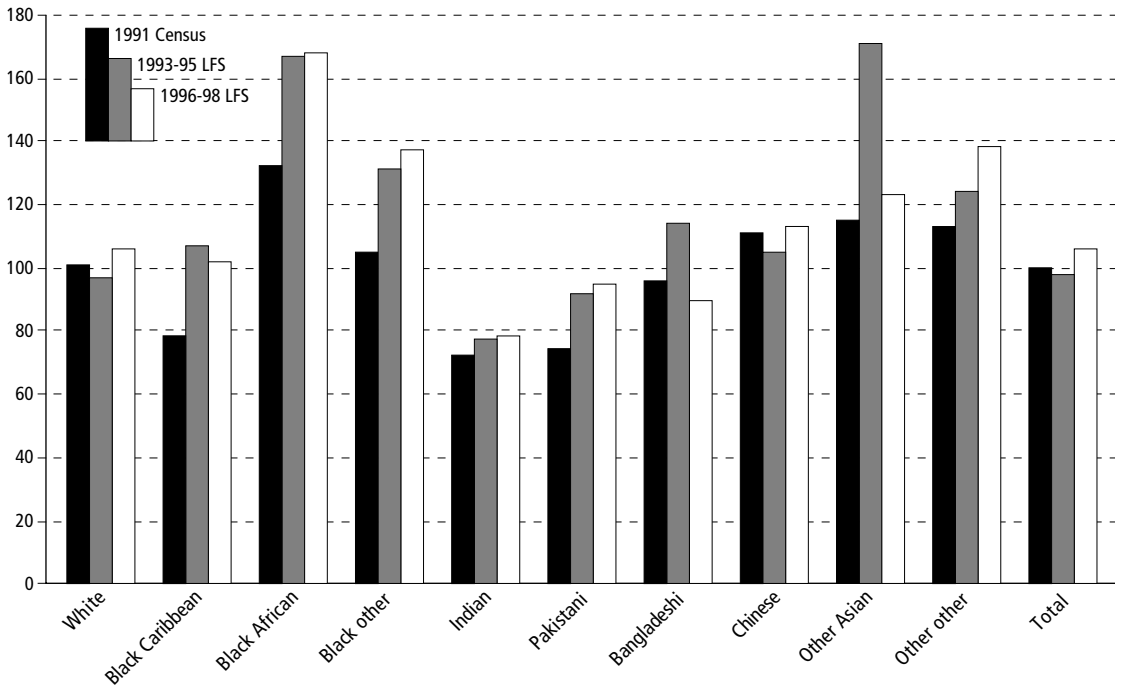
8.3 Combining data sources together

8.3.1 Administrative data and the Census

In response to the demand from local authority planning departments, however, some organisations have produced population projections by ethnic group. They have been obliged to use imperfect data to derive the internal migration component of the projections. Three examples of such sets of projections are those produced for Bradford Council (Simpson, 2000), Oldham Metropolitan Borough Council (OMBC, 1997) and the LRC (Storkey, 2000b). In all three cases, to forecast the population, census-based estimates of migration for the different ethnic groups were constrained to the official NHSCR-derived total inflows and outflows for all population groups in the early 1990s. The Bradford model incorporated an additional refinement. This involved adjusting the migration figures to take account of the differential rates of change in the numbers of people with South Asian and non-South Asian names on the Bradford electoral register. Unfortunately, however, with the exception of the Chinese,

Figure 2

Comparison of 1991 Census and Labour Force Survey migration data: migrants within regions



members of other ethnic groups are not easily distinguishable by name. Thus, the analysis of names on registers is of limited value in areas with mixed ethnic minority populations.

8.3.2 Linked patient registration and Longitudinal Study data

In the preparation of internal migration estimates, ONS compiles a patient register for the whole of England and Wales. While the register does not currently include information on patients' ethnic group, it may be possible to link register data – via the patients' NHS numbers – to the records for the 1 per cent of individuals in the Longitudinal Study, so accessing information on their ethnic group.

An enhanced LS could be used to examine trends in the migration of the minority ethnic populations in the years leading up to the 2001 Census. A crucial advantage of this would be the size of the sample. It would be more than three times that of the current LFS sample.

However, there would be problems in linking the two data sets.

- The first and most important problem is that adding the patient registration data to the LS would require both the permission of the Department of Health and an assessment of the feasibility of the work by the Migration Statistics Unit, and the LS team. Careful consideration would need to be given to assessing the legal implications of linkage, as well as evaluating the likely costs against benefits.

- Secondly, the patient register estimates of migration have some limitations. First, postcodes of origin and/or destination are missing or invalid for a very small proportion (for example, 0.5 per cent in 1999) of patients on the health authority registers and secondly, certain groups of migrants, who, for one reason or another, were not registered in any one year, are missing. For this reason, data are combined with information recorded at NHSCR. There is also evidence that the patient registration data underestimate young adult male migrants. This problem will be investigated further when the 2001 Census results become available.

- Ethnic group (and country of birth) information is currently not available for people who have joined the LS since 1991. These comprise new births, new immigrants with LS birth dates and LS members who have re-entered the country but were not present at the Census. Thus, migration estimates for the late 1990s for people of known ethnic origin will be biased towards the mobility experience of long-term residents and will not incorporate the, potentially different, migration patterns of the large number of recent arrivals in the country. It would be necessary therefore to use proxy data to an assign ethnic group to recent immigrants.

Table 3 shows numbers of male one-year migrants in the LS for 1991 and some ratios of the observed to the expected numbers calculated from the 100 per cent Census figures. Most of the ratios are close to 1.0, but a few are significantly lower than 1.0 at the 5 per cent level (Table 3). Part of the explanation for the low ratios may be that people who moved in the year before Census are less likely to have been traced by the LS than people who did not move. This possibility is being investigated.

8.4 Making an estimate of internal migration of ethnic minorities

Measuring migration within the United Kingdom is not straightforward, as there is no system to record the movement of people. For the minority ethnic populations of England and Wales, the 1991 Census provides the most comprehensive existing data on internal migration.

When the 2001 Census results are published, a second set of migration statistics will become available. The combination of these results with those from the 1991 Census will provide estimates of migration that are adequate for population projections for regions and counties and, for the larger minority ethnic populations, for local authority districts or unitary authorities. Methods will be needed to develop a 'mapping' from the ethnic group categories used in the 1991 Census to those used in the 2001 Census, which could be done using census country of birth data for different age groups or by comparing LFS results for the years covering the changeover from the 1991 to the 2001 question.

If estimates on an annual basis are required, then the information derived from the Census, together with an estimate of 'White' internal migration, would need to be constrained on an annual basis to the estimates of internal migration of the total population derived from re-registrations with GPs.

The Labour Force Survey, the other existing data source on internal migration by ethnic group, can provide a useful insight into migration patterns over time. But, because of the small sample size, conclusions about migration trends cannot be drawn from these data. However, if the two censuses reveal substantial differences in

Table 3

Male migrants in the Longitudinal Study and ratios of observed to expected migrants, 1991 by ethnic group

Migrants/Ratios	Total	White	Black Caribbean	Black African	Black other	Indian	Pakistani	Bangla-deshi	Chinese	Other Asian	Other other	All ethnic minorities
Within districts												
All ages	13857	12817	125	109	70	256	164	83	51	70	112	1040
1-4	1380	1277	15	9	20	18	11	5	5	7	13	103
5-15	1725	1554	12	11	17	37	35	26	3	6	25	171
16-19	698	637	3	7	5	10	10	8	2	6	9	61
20-29	4691	4415	34	40	15	57	39	16	16	25	34	276
30-44	3191	2913	36	36	11	83	39	13	17	18	25	278
45-64	1350	1232	16	6	2	41	24	13	5	8	3	118
65+	822	789	9	0	0	10	6	2	3	0	3	33
Between districts within counties												
All ages	3096	2865	27	44	8	51	27	21	8	18	27	231
1-4	185	171	2	5	0	1	2	2	0	1	1	14
5-15	282	254	2	1	2	8	6	2	0	3	4	29
16-19	132	118	2	2	0	4	2	3	0	0	1	13
20-29	1268	1185	16	18	3	13	8	5	3	6	11	83
30-44	790	729	3	13	2	13	5	3	5	8	9	61
45-64	317	290	2	5	1	11	4	4	0	0	0	27
65+	122	118	0	0	0	1	0	2	0	0	1	4
Between counties within regions												
All ages	2616	2372	37	34	12	48	28	17	20	17	31	244
1-4	170	156	3	1	0	2	1	2	1	1	3	14
5-15	217	192	4	1	4	6	3	0	0	3	3	25
16-19	88	73	4	4	2	0	1	2	3	0	0	15
20-29	1040	942	17	18	3	19	7	9	8	5	12	98
30-44	739	670	7	10	3	14	12	2	5	5	11	69
45-64	270	248	2	0	0	7	4	2	3	3	1	22
65+	92	91	0	0	0	0	0	0	0	0	1	1
Between regions												
All ages	3031	2867	13	7	6	51	23	11	24	9	20	164
1-4	207	195	0	0	1	6	2	0	0	1	2	12
5-15	340	326	1	1	1	4	2	2	1	0	2	15
16-19	182	172	1	0	0	2	1	1	3	0	2	9
20-29	1099	1038	8	3	2	18	6	3	10	2	9	61
30-44	754	699	1	1	1	19	12	2	8	6	5	55
45-64	317	307	2	1	1	2	0	3	1	0	0	10
65+	132	130	0	1	0	0	0	0	1	0	0	2
All migrants	22600	20921	202	194	96	406	242	132	103	114	190	1679
Ratios of observed to expected[1] migrants												
w'in Ds	0.99	0.98	1.05	1	0.17	1.10	1.05	0.84*	1.19	0.90	0.95	1.06
1.02												
b D w'in Cs	0.93*	0.94*	0.95	1.01	0.61	0.75*	1.16	0.89	0.40*	0.73	0.87	0.84*
b Cs win Rs	0.93*	0.92*	1.24	0.85	0.94	0.84	1.22	0.97	0.93	0.67	1.01	0.95
Betw' Rs	1.00	1.00	0.95	0.53	0.67	0.94	0.84	1.00	1.27	0.62	0.84	0.88
All migrants	0.97*	0.97*	1.06	1.03	0.97	0.96	0.90	1.08	0.88*	0.83	1.00	0.97

[1] *Calculated from 100 per cent Census data*
* *Ratio significantly different from 1.00 at 5 per cent level*
Source: ONS Longitudinal Study and 1991 Census

migration rates, the LFS statistics may help to establish the timing of the changes and factor them into the derived annual estimates that are constrained to the published migration data. If necessary, linking patient re-registrations to the records for the individuals in the Longitudinal Study could yield further imperfect, but useful, information on changing migration patterns over time.

8.5 Projecting internal migration by ethnic group

The current subnational population projections model used by ONS to produce population projections of the total population for all local and heath authorities, has the facility to project the internal migration of up to ten different population types. As part of this model, it is also possible to transfer between the appropriate population types. The model has two processes: firstly, a projected outflow is established for each area and population type and secondly this outflow is distributed to form the inflows for each area and population type. These calculations are carried out by gender and age group. It is therefore possible to have different migration propensities and patterns for the different ethnic groups.

It would in theory be possible to utilise the functions in the ONS model to project each component of population change, eg fertility and mortality by ethnic group. This would produce a set of population projections by ethnic group that are consistent with the national and subnational population projections currently produced by the Government Actuary's Department (GAD) and ONS. While, this is outside the scope of this paper on internal migration, it maybe useful to those working on methodology.

References

Champion T (1996) Internal migration and ethnicity in Britain, in Ratcliffe P (ed), *Ethnicity in the 1991 Census, volume 3: social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration, ONS*, HMSO: London, pp 135–73.

Chappell R, Vickers L and Evans, H (2000) The use of patient registers to estimate migration. *Population Trends* 101, pp 19–24.

Oldham Metropolitan Borough Council (1997) *Population projections for Oldham*, OMBC, Chief Executive's Department, Policy Unit: Oldham.

Rees P H and Duke-Williams O (1997) Methods for estimating missing data on migrants in the 1991 British Census. *International Journal of Population Geography*, vol 3 no. 4, pp 323–68.

Scott A and Kilbey T (1999) Can patient registers give an improved measure of internal migration in England and Wales? *Population Trends* 96, pp 44–55.

Simpson L (2000) Migration within the UK, in *Population change in the Bradford District: age sex and ethnic group*, City of Bradford Metropolitan District Council, Policy and Research Unit: Bradford, pp 24-30.

Storkey M (2000a) Migration,. in *Without prejudice: exploring ethnic differences in London*, London Research Centre: London, pp 35–51.

Storkey M (2000b) *Methodology of the ethnic group projections 1999 Round*, London Research Centre: London.

This is a blank page

Chapter 9

Mortality data on migrant groups living in England and Wales: issues of adequacy and of interpretation of death rates

Seeromanie Harding and Rasaratnam Balarajan

9.1 Background

There is a paucity of mortality data for minority ethnic populations living in England and Wales and what is available is subject to many limitations. In most studies, birthplace is used as a proxy for ethnicity because of the lack of relevant data on ethnic origin, thereby excluding second or later generations born in the UK. There are two main national data sources used to study the death rates of migrants living in England and Wales. There are some local studies but most of these are ad hoc or designed to meet a specific epidemiological need. We briefly discuss the national data sources and their limitations before highlighting the issues that may need methodological caution if these data are used for projecting populations by ethnic group.

9.2 Data

A summary of the data sources discussed below, and what they contain, is shown in Table 1.

9.2.1 Cross-sectional national data routinely collected

As stated above, most studies in migrant mortality use country of birth as a proxy for ethnicity. The first of these decennial analyses was undertaken using data around the 1971 Census, 1970–78 (Marmot et al, 1984), and was repeated using data around the 1981 Census, 1979–83 (Balarajan et al, 1984), and the 1991 Census, 1991–93 (Harding and Maxwell, 1997; Wild and McKeigue, 1997). Table 2 shows the standardised mortality ratios and deaths (SMR) for the 3 periods. The findings from these analyses remain the

Table 1

Data sources which can be used to calculate mortality rates of migrant populations in England and Wales

Sources	Age in single years	Age in 5 year age groups	Country of birth	Ethnic origin	Country of birth of grandparents or parents	Ethnic origin of parents	Coverage	Quality of data
Unlinked								
Death registration	Yes	Yes	Yes	No	No	No	100%	Country of birth reported by relative of deceased maybe different from that reported in census
Census	Yes*	Yes	Yes	1991/2001	Question asked in 1971 Census	No^	100%	1. Under-enumeration of some groups in specific geographical areas 2. May contain migrants who return to home countries, leading to inflation of denominators.
Linked								
ONS Longitudinal Study	Yes	Yes	Yes	1971 (South Asians)~ 1991/2001	Data in non-members' files, could be linked to members' data	Data in non-members' files, could be linked to members' data	1%	1. Loss to follow-up at older ages can be up to 30% for some groups 2. Small sample size

** Not in published tables but date of birth is recorded at census so could be derived from original census data.*

~ Question not asked in 1981 and 1991 Censuses but if the person was usually resident in the household it could be derived using relationship to head of household.

^ Question not asked but if the person was usually resident in the household, ethnic origin could be derived using relationship to head of household.

~ Special exercise to identify South Asians by names

most important for monitoring death rates of migrant groups over the last two decades. Significant though these studies are, they are all limited by the use of country of birth as a proxy measure for ethnicity. This would have been relatively unimportant in 1971 but by 1991, the use of this proxy measure meant that significant numbers of second and later generation individuals would have been excluded for the analyses. About one half of those who reported a Black Caribbean ethnic group at the 1991 Census were born in the UK.

Death rates are calculated using the age distribution of the population by country of birth from the censuses and the age distribution of the deaths by country of birth from death registrations in the years around the censuses. The small number of deaths among those aged under 20 years is usually a problem even if the data are aggregated across three to five years. Most epidemiological studies have thus reported age adjusted mortality risks for migrants aged 20–69 years and 70 years and over. Table 3 shows numbers of deaths for the years 1988–92 by country of birth and five-year age groups for men and women in selected migrant groups living in England and Wales.

9.2.2 ONS Longitudinal Study (LS) data

Information from death certificates is linked to the records of the study members in the LS, a record linkage study based on 1 per cent of the population of England and Wales (Hattersley and Creeser, 1995). With over 25 years of follow-up for the 1971 cohort (people present in the 1971 Census), the value of this data set is now being realised. With only 1 per cent of the population involved, the number of people – and of their deaths – in the migrant groups is small, but the samples are representative of the large migrant groups living in England and Wales. Cause-specific patterns of mortality for those from the Indian subcontinent and from the Caribbean Commonwealth in the 1971 cohort are similar to

Table 2

Standardised Mortality Ratios (SMR) of migrants in England and Wales in the 1970s, 1980s and 1990s, aged 20-69✦

	1970-72^ Males	Females	1979-83* Males	Females	1988-92+ Males	Females
Scotland						
SMR	112	107	118	119	129	132
Deaths	8561	4229	14478	7628	12785	7562
Ireland						
SMR	121	116	129	120	140	119
Deaths	10547	5590	18961	10403	17015	9298
Indian subcontinent						
SMR	99	111	107	106	103	96
Deaths	3007	1442	7566	3427	9219	4329
Caribbean Commonwealth						
SMR	95	131	79	104	83	96
Deaths	1326	878	3146	1922	4012	2401
East Africa~						
SMR			110	109	103	95
Deaths			791	390	1395	679
Other Commonwealth Africa~						
SMR			83	90	102	110
Deaths			451	189	671	331

✦ the standard rates for SMR for 1970-72 are age and sex specific rates for all England and Wales in 1970-72
✦ the standard rates for SMR for 1979-83 are age and sex specific rates for all England and Wales in 1979-83
✦ the standard rates for SMR for 1989-92 are age and sex specific rates for all England and Wales in 1989-92
~ Data not available for 1970-72
^ Data for 1970-72 from Marmot M G, Adelstein A M, and Bulusu L. Immigrant Mortality in England and Wales 1970-78 OPCS Studies of Medical and Population Subjects No 47. London: HMSO,1984.
**Data for 1979-83 from Balarajan R and Bulusu L. Mortality among immigrants in England and Wales, 1979-83. Mortality and Geography - a review in the mid-1980s, England and Wales DS 9: 104-121. London; HMSO.1991*
+Data for 1989-92, unpublished results derived by authors

those found in the national (100 per cent) cross-sectional data, reflecting that the data are broadly representative of the populations.

Table 4 shows SMRs and number of deaths for persons aged 25–64 in 1971 and followed up for about 19 years. Identification of South Asians in this cohort was improved by the analysis of names (Stuart, 1990), thereby excluding children born in the Indian subcontinent to parents who were British expatriates. Another advantage of these data is that the deaths are unlikely to be misclassified, as they are classified by the country of *birth reported in the census*, whereas in the unlinked data, deaths are classified by the country of birth reported *at death registration* – and the populations used in the denominators are by country of birth from the census. The country of birth reported by relatives of the deceased is not always congruent with that reported in the census (see Section 9.3.1).

In addition to the question on country of birth of those usually resident in the 1971 Census, a question on country of birth of parents was also asked. In the LS, data on the non-LS members of the LS members' households are available, making it possible to classify people by parental birthplace. Generational comparisons in mortality are therefore possible in principle, but in reality this exercise is limited because of small sample sizes.

Table 3

Number of deaths in England and Wales by country of birth and age in the years 1988-92

	Males																	
Country of birth	**0-4**	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
England and Wales	18402	1666	1694	5983	8633	8691	8883	11464	18026	25888	40188	66698	114433	180459	209172	252590	222953	187128
Scotland	10	8	14	81	160	199	268	349	555	838	1192	1772	2782	4670	5414	5026	3775	2871
All Ireland	22	8	9	43	111	151	197	353	629	1009	1820	2734	4116	5895	6260	6027	3752	2328
India	2	2	3	13	30	53	95	180	273	450	721	1041	1363	1590	1650	1532	1263	1027
Pakistan	11	10	17	17	30	32	64	111	81	139	295	402	488	435	229	141	88	119
Bangladesh	3	8	6	12	9	7	8	17	17	64	235	295	254	149	65	40	10	8
Sri Lanka	1	1	0	2	9	6	10	17	17	27	45	42	68	50	69	77	59	44
Caribbean Commonwealth	1	0	0	1	13	30	93	98	91	258	544	840	1024	1021	829	598	259	140
East Africa	4	4	0	24	46	104	132	145	163	144	160	194	195	112	81	44	28	16
Other Commonwealth Africa	6	5	1	13	29	39	50	54	45	54	85	103	105	107	85	55	36	25
	Females																	
Country of birth	**0-4**	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
England and Wales	13574	1110	1056	2365	3097	3605	4702	7032	11978	16825	24729	40526	70732	117518	155465	230155	284022	460465
Scotland	6	7	8	34	41	68	104	201	268	487	645	952	1659	3137	3528	4096	4323	6654
All Ireland	9	3	3	15	28	35	74	160	345	596	858	1495	2371	3336	4723	5317	4758	5983
India	1	1	2	7	21	37	53	81	140	205	350	489	781	1059	1252	1532	1479	1710
Pakistan	3	4	6	18	30	25	49	44	57	57	109	165	157	121	97	111	88	108
Bangladesh	2	2	5	4	12	7	8	11	14	21	30	29	36	14	15	7	9	10
Sri Lanka	1	0	1	1	2	3	7	3	7	12	12	16	27	28	55	68	62	105
Caribbean Commonwealth	1	0	0	2	2	14	61	85	107	219	340	489	559	525	459	362	291	348
East Africa	6	1	2	8	35	43	58	79	69	57	107	96	72	63	56	50	21	21
Other Commonwealth Africa	4	2	2	10	10	23	34	26	50	48	45	37	25	33	26	30	29	15

Table 4

Standardised Mortality Ratios* (SMR), 1971-89, for migrants born in the Indian subcontinent, Caribbean Commonwealth, Republic of Ireland, Northern Ireland and Scotland, aged 25-64. England and Wales. ONS Longitudinal Study.

	Males	Females
Indian subcontinent		
SMR	103	75
Deaths	162	28
Caribbean Commonwealth		
SMR	63	114
Deaths	70	56
Scotland		
SMR	121	110
Deaths	261	146
Republic of Ireland		
SMR	112	103
Deaths	252	151
Northern Ireland		
SMR	119	112
Deaths	380	46

** The standard rates for SMR are age, year and sex specific rates for all Longitudinal Study members*

Table 5

Number of persons and deaths amongst first and second generation ethnic groups aged under 65 years in the 1991 Census and living in England and Wales. ONS Longitudinal Study, 1991 Cohort followed up 1991-97

Ethnic origin at the 1991 Census	Sample	Deaths
Whites born in the United Kingdom	**394220**	**7510**
South Asian groups		
First generation Indian	6968	138
Second generation Indian	3801	11
First generation Pakistani	3557	51
Second generation Pakistani	2847	11
First generation Bangladeshi	1866	20
Second generation Bangladeshi	613	1
First generation South Asians	*12391*	*209*
Second generation South Asians	*7261*	*23*
African groups		
First generation Black Caribbean	2306	76
Second generation Black Caribbean	2728	13
First generation Black African	1562	15
Second generation Black African	794	5
Chinese		
First generation Chinese	1213	13
Second generation Chinese	438	0

Ethnic group was recorded in the 1991 Census – the first census to do so – and is linked to the records of LS members. Table 5 shows the number of LS members in each ethnic group present in 1991 and the number of deaths for the period of follow-up, 1991–97. Longer follow-up of this cohort will offer opportunities, hitherto unavailable, for measurement of death rates by ethnic group, rather than by country of birth. Small sample sizes of LS members, however, make it unusable at a local area level for most groups. A category for 'Irish' is not shown in this table because it was not included in the 1991 Census classification.

9.3 Methodological cautions

9.3.1 Misclassification at census and at death registration

The first decennial analyses of ethnic differences in mortality (Marmot et al, 1984) identified problems of ethnic misclassification – which remains a relevant consideration. The misclassification at death of older people born in what is now Pakistan, but what was then India, resulted in unreliable estimates of mortality for those of Pakistani origin. In the last set of decennial analyses of death rates of migrants around the years of the 1991 Census, the LS was used to compare country of birth reported at death with country of birth reported in the census for the same people (Harding and Maxwell, 1997). The problem was not as great as in the previous analyses using data in the 1970s, but a significant minority of those born in India and Pakistan were still being inconsistently classified at the older ages. There is a high degree of congruence between Indian, Pakistani and Bangladeshi reportage, and country of birth reportage in the 1991 Census for those aged under 69 years (Office of Population Censuses and Surveys and General Register Office for Scotland, 1993), which suggest

that the census classification, rather than that at death registration, is the correct designation. At older ages, however, there is clearly still the potential for inconsistent classification in the numerators and denominators of the death rates.

9.3.2 Heterogeneity within country of birth groupings

A related point refers to the use of aggregated country of birth groupings such as 'Indian subcontinent' (usually referred to as 'South Asian' or 'Asian'). There is evidence to suggest that this term is too broad to be meaningful because of the cultural and economic diversity within the 'South Asian' grouping (Harding and Maxwell, 1997; Balarajan and Raleigh, 1997; National Institute for Ethnic Studies in Health and Social Policy, 1997; Nazroo, 1997; Bhopal et al, 1999; Harding, 2000). Mortality of these groups also differs significantly, partly because of differences in poverty, Bangladeshi men having higher mortality rates than Indian men (Table 6) (Harding, 2000).

Other work has also echoed a similar view, even adding further caution by highlighting that there are differences *within* the separate countries of the Indian subcontinent. The analysis of names on death certificates was used in the 1980s to identify South Asians from different regions of the subcontinent (Balarajan et al, 1984). This analysis involved a manual inspection of all death records for three years and provided valuable information on the heterogeneity of cause-specific mortality risks among, for example, Punjabis, Gujeratis, Muslims and Hindus.

It is possible that Nam Pehchan, the computer program developed by Bradford District Council, could be used in the future (see Chapter 1 in this volume) to identify different South Asian subgroups. The findings of exploratory work using death registrations for one year in Bradford showed that prediction could be more than 90 per cent correct when results based on visual inspection of the records were compared with computerised classification of different groups of South Asians (Harding at al, 1999). This approach would seem to be worth investigating further as it may provide the opportunity to classify South Asians into meaningful categories at a national level, assuming that the hurdles relating to confidentiality can be overcome.

The importance of meaningful classifications cannot be emphasised enough if projections of these populations are to be used for the planning and provision of clinical and supportive services. For example, Bangladeshis (a quarter of who live in Tower Hamlets) are younger relative to Indians, but stroke mortality is considerably higher – even among those aged under 65 years (Balarajan and Raleigh, 1997; Harding, 2000). If supportive care for stroke patients in Tower Hamlets were planned on projected populations of over 65 year olds 'South Asians', there would be an underestimation of need within the Bangladeshi community.

Similarly, Caribbean-born people are assumed to be of African origin but a small, but significant, proportion is of Indian ancestry. There are also significant inter-island differences in history and culture and a classification that combines, for example, the Nevisians with the Jamaicans, though convenient, becomes problematic. The commonly-used country of birth grouping 'Born in African Commonwealth countries' is another example of assuming homogeneity among people as diverse as those from Somalia, Mozambique, Kenya and Tanzania with vast cultural diversity. Migrants in the UK from eastern African

Table 6

Mortality in England and Wales, 1989-92, of migrants born in India, Pakistan and Bangladesh, aged 20-69✦

	Males	Females
Indian subcontinent~		
SMR	103	96
Deaths	9219	4329
India		
SMR	104	103
Deaths	5796	3216
Pakistan		
SMR	95	88
Deaths	2077	814
Bangladesh		
SMR	118	71
Deaths	1055	182
Sri Lanka		
SMR	82	61
Deaths	291	117

~ Includes those born in India, Pakistan, Bangladesh and Sri Lanka

✦ the standard rates for Standardised Mortality Ratios (SMR) are age and sex specific rates for all England and Wales in 1989-92

countries such as Tanzania are mainly of Indian Gujerati origin, and those from Mozambique and Angola are mainly of African ancestry but the former are French speaking and the latter, Portuguese speaking, in addition to their varied tribal histories.

9.3.3 Children born in the Indian subcontinent whose parents were British expatriates

As alluded to above, the use of country of birth includes the children of British expatriates living abroad, most of whom returned to the UK after political independence of the colonies. Marmot et al (1984) estimated that up to 20 per cent of those born in the Indian subcontinent who were aged 50 years and over in 1971 could have been of British ancestry. There remains much controversy over the extent to which this factor affects the mortality rates at older ages, as there is little data from which to draw firm inferences. Arguably we should see mortality rates intermediate to those of Indian origin and those of British origin, given that exposures in childhood (in this case in the Indian subcontinent) are important determinants of mortality risk in adult life. Some preliminary evidence from the LS provides confirmation. Table 7 shows the risk of mortality of India-born Indians and India-born Whites relative to UK-born Whites using hazard ratios derived from Cox regression models. The hazard ratios show that, amongst those aged 65 and over, mortality of Whites born in India is higher than that of Indians born in India, though lower than that of UK-born Whites.

Table 7

Mortality in England and Wales, 1991-98, of those born abroad by ethnicity reported at the 1991 Census and aged 65 and over at the census, ONS Longitudinal Study

	Hazard Ratio	95% CI
Whites born in England and Wales	**1.00**	
Born in India		
White	0.89	0.74-1.08
Indian	0.32	0.20-0.52

9.3.4 Return migration of older migrants

All-cause mortality of Caribbean-born people is low relative to the UK average, attributed mainly to their low mortality from coronary heart disease. Paradoxically, their mortality from hypertension and diabetes, risk factors for coronary heart disease, and from stroke, which shares the same risk factors, is very high, but because coronary heart disease accounts for a large proportion of the deaths, their overall mortality rate is low. An unexplained element, however, is the impact of return migration on the mortality rates, particularly at older ages.

From the LS, we know that loss to follow-up of study members is greatest (up to 30 per cent) at older ages. In the LS, loss to follow-up is assumed if study members are not found at subsequent censuses or by registration of a vital event. There are documented life histories of return migrants to the Caribbean by Chamberlain (1997) and Byron (1994), mainly consisting of those who return on retirement, and to a lesser extent those who return prior to retirement and children of school age who return with their parents. 'Short stay' migrants tend to return within 14 years whereas 'long stay' migrants tend to return after 30 years.

There is evidence of elderly people returning to the Caribbean to take advantage of the lower cost of living and to get more from their UK pensions. Some migrants may even have multiple residences (Grasmuck and Pessar, 1991; Byron and Condon, 1996). There is also considerable return migration amongst those from Indian subcontinent and from the Irish Republic at older ages. However, there are others, such as Anwar (1979), who argue that while most migrants had a return agenda in their migration plan (in this example Pakistan), few would realise it. Byron and Condon (1996) also point out that a gender-balanced Caribbean migrant population resulted in marriage and family formation, which strengthened links with Britain and complicated the return issue.

It is important to consider how return migration can affect the death rates. Return migration could lead to an underestimation of the death rates shown in Table 8, particularly at older ages. If people were counted at the census but left the country after the census, this would inflate the denominators and lead to an underestimation of death rates. Secondly, if return migration is selective in terms of health, in that the less healthy return to die in their home countries, then this could depress the death rates, the effect being

Table 8

All-cause age specific death rates per 100,000, 1988-92, by country of birth, sex and age, England and Wales

Country of birth	Males 20-29	30-39	40-49	50-59	60-69	70+	Females 20-29	30-39	40-49	50-59	60-69	70+
England and Wales	91	118	261	811	2469	8365	34	67	171	492	1415	6515
Scotland	123	172	363	1078	3100	8619	39	92	218	676	1902	6958
All Ireland	110	217	453	1223	3207	8672	24	85	243	614	1621	6470
India	80	123	357	872	2328	7704	41	56	159	482	1509	6375
Pakistan	62	108	281	795	2148	4877	44	60	138	437	1057	3499
Bangladesh	35	55	453	1143	2275	4695	36	41	113	273	1110	2336
Sri Lanka	67	102	202	651	1908	7697	34	43	102	274	826	5561
Caribbean Commonwealth	135	153	348	718	1709	5202	37	85	220	457	1178	4315
East Africa	111	134	303	764	1810	4152	57	73	143	480	876	2922
Other Commonwealth Africa	115	122	245	726	2354	6515	43	70	221	463	1154	3617

larger if they were included in the census. On the other hand, if people returning are healthier than those remaining, as may be the case for pre-retirement ages, then this could result in higher than expected death rates.

In general, there are very few empirical studies that look at the health of return migrants and our understanding of the nature of return migration – and its impact on the death rates of migrants in the host country – is limited. This is a problem common to migrant studies all over the world. In Sweden, researchers tried adjusting the denominators (which they felt could have been inflated because of return migration) of the estimated death rates by using income received in the last year to confirm presence of the individual in Sweden at the time of exposure to risk of dying (Weitoft et al, 1999). This produced significant shifts in the death rates. In the UK, there are no comparable methods to adjust the denominators.

9.3.5 Generational differences in mortality

This is an area of work of increasing importance, given the large numbers of second and later generations born in the UK. Most of what we know about generational differences in mortality comes from the work on Irish migrants living here (Abbotts et al, 1998; Harding and Balarajan, 1996; Harding and Balarajan, 2001). We have argued elsewhere, in relation to Irish migrants, that the evolution of ethnic identity through generations is influenced by interaction with the social, economic and political environment of the host community (Harding and Balarajan, 2001). This argument is relevant to all ethnic groups and the experience of children and grandchildren of the Irish is worthy of some discussion, as, although there are dissimilarities, there are also similarities in experiences, and it may be a useful starting point in making assumptions about the experience of second and later generations of other migrant groups in the UK.

As the largest ethnic group, the estimated number of second generation Irish exceeds two million. An Irish category has been included for the ethnic group question in the 2001 Censuses for England and Wales, and Scotland. Work so far has used country of birth of parents and grandparents in the 1971 Census to identify second and third generations of Irish people. Mortality has risen progressively across the three generations – despite upward social mobility (Table 9) (Harding and Balarajan, 2001). Exploratory work on generational differences in the mortality of Indians, Pakistanis and Bangladeshis shows that mortality of Pakistanis born here is considerably higher than that of those born in Pakistan (Table 10). Most of this excess mortality was due to the high death rates of those aged under 16 years. The numbers of deaths were small in this study, but poorer health of Pakistani children, linked to the high incidence of congenital abnormalities, compared with Indian children, has been confirmed by other studies (Bundey et al, 1991; Powell et al, 1995; Powell et al, 1994). Very high mortality rates for infants born to Pakistani-born

mothers are also well documented (Table 11) (Davey Smith et al, 2000).

A consistent feature of a recent review of the literature on generational differences in health (Arai and Harding, 2002) was the poorer health of the children of migrants relative to the health of first generation migrants. There is abundant evidence from US studies which have shown that infant mortality is lower amongst infants born to foreign-born Black mothers than amongst those born to US-born Black women (Cabral et al, 1990; Hummer et al, 1999; Pallotto et al, 2000). Infant health is significantly related to health in later life, and the logical extension is that the mortality experience of these later generations are likely to be different from that of the first generation.

Fang et al (1996) showed a complicated crossover occurring across generations among those under 65 years; lowest mortality in Caribbean-born men and women, and highest in US-born Blacks with the US-born White population occupying an intermediate position. Information on length of residence was not available, but the authors suggest that most of the migrant Black Caribbeans had been in the US for 13 years. We return to the issue of the relationship between mortality and the length of residence in the host country later. The important point here is that US-born Blacks and Caribbean-born Blacks share a common West African ancestry but show different mortality patterns, no doubt related to different exposures in the different environmental settings.

A pertinent point that relates to distinguishing second or later generations is that of changing identities across generations and the resulting complications in classifying later generations. In a recent study, there was a dramatic difference in reported rates of ill health between Black Africans born in the UK and those born in Africa (Harding and Balarajan, 2000). Whereas the rate for Black Africans born in Africa was lower compared with that for the White group, the rate for Black Africans born in the UK was higher. Further examination of the data revealed that a small proportion of those who described themselves as migrant 'Black Africans' had been born in the Caribbean. These people had not identified with the 'Black Caribbean' option. This makes it difficult to assess how many of those who were born in the UK and described themselves as 'Black Africans' were actually of Black Caribbean parentage. The growing numbers

Table 9

Number of persons, deaths and hazard ratios for all cause-mortality 1971-97, for first, second and third generations of Irish people living in England and Wales. 1971 ONS Longitudinal Study cohort.

	Sample	Hazard ratio	95% CI	Deaths
Males				
Other	17153	1.00		24829
First generation born Northern Ireland	735	1.27*	(1.11-1.47)	195
First generation born Irish Republic	2052	1.13*	(1.04-1.23)	556
Second generation Irish	4564	1.29*	(1.18-1.41)	503
Third generation Irish	2513	1.31*	(1.06-1.63)	86
Females				
Other	166861	1.00		16110
First generation born Northern Ireland	726	1.25*	(1.04-1.49)	125
First generation born Irish Republic	2224	1.14*	(1.03-1.26)	386
Second generation Irish	4420	1.21*	(1.09-1.36)	314
Third generation Irish	2384	1.55*	(1.17-2.05)	50

Source - Harding S and Balarajan R. Mortality of third generation Irish living in England and Wales.BMJ (2001).

Table 10

All-cause mortality of first and second generation ethnic groups aged under 65 at the 1991 Census and living in England and Wales. ONS Longitudinal Study, 1991 Cohort followed up 1991-97.

Ethnic origin at the 1991 Census	Hazard ratio	95% Confidence Interval	Deaths
Whites born in the United Kingdom	**1.00**		**7510**
South Asian groups			
First generation Indian	0.84*	(0.71-0.99)	138
Second generation Indian	1.28	(0.71-2.31)	11
First generation Pakistani	0.77	(0.59-1.02)	51
Second generation Pakistani	2.48*	(1.37-4.48)	11
First generation Bangladeshi	0.64	(0.41-1.00)	20
Second generation Bangladeshi	0.88	(0.12-6.26)	1
First generation South Asians	*0.80**	*(0.70-0.92)*	*209*
Second generation South Asians	*1.62**	*(1.07-2.45)*	*23*
African groups			
First generation Black Caribbean	0.80	(0.64-1.00)	76
Second generation Black Caribbean	1.15	(0.67-1.99)	13
First generation Black African	0.70	(0.42-1.16)	15
Second generation Black African	1.85	(0.77-4.46)	5
Chinese~			
First generation Chinese	0.57*	(0.33-0.98)	13
Second generation Chinese	-	-	0

Table 11

Infant mortality rates, 1996, by mother's country of birth, England and Wales

Country of birth of mothers	Infant Mortality Rate per 1000 births
UK	5.8
India	5.4
Pakistan	10.1
Bangladesh	6.3
Caribbean Commonwealth	8.4
East Africa	6.1

of interethnic unions and increasing generational distance between British-born Black Caribbeans and the Caribbean 'homeland' have implications for the ethnic identity of second and later generations.

9.3.6 Changes in mortality rates with increased length of residence in the United Kingdom

Few studies have been able to examine the health of migrants in relation to their length of residence in the host country. From the adaptation model, we expect migrants to take on the health patterns of their host countries after the initial effects of migration, such as the healthy migrant effect, or the stress of migration, have worn off (Haenszel and Kurihara, 1968; McMichael and Bonett, 1981). Work in the 1960s on the Japanese in America was the first to demonstrate this trend (Haenszel and Kurihara, 1968). A similar trend for South Asians living in Canada has been suggested (Sheth et al, 1999). However, as studies of Turkish migrants in Germany have shown, mortality can remain lower than that of the indigenous population for more than 15 years and persist into the second generation (Razum et al, 1998; Razum et al, 2000). The 'healthy migrant effect' refers to the methodological problem that confronts migrant studies due to self-section at the time of migration; healthy and younger individuals tend to migrate while those with serious health problems or are older tend not to. Voluntary migrants are also likely to be healthier than involuntary migrants. This effect should diminish with time, but could last 20 years or more (Kliewer, 1992). It is a potential effect that is still not properly understood or quantified. The historical time of migration also needs to be taken into account; later migrants may differ in educational and other demographic characteristics.

Preliminary findings from a study of South Asians in England and Wales suggest that all-cause mortality increases with length of residence (Hardingand Balarajan, 2002). Long distance migration is subject to fairly stringent criteria and migrants tend to be healthier than the population they have left. We would expect this health advantage, however, to wear off as their exposure to changes in diets, exercise levels, alcohol and cigarette consumption increases. In addition migration to the UK after the 1960s was more skilled, implying better health and higher socio-economic status than of earlier arrivals (Heath and Ridge, 1983).

Table 10 shows that mortality is consistently lower for the first generation migrants compared with that of the White population born in the UK. Although, most of these migrants arrived before 1970, the low mortality is undoubtedly influenced by the 'healthy migrant effect'. Year of immigration was not recorded in the 1991 Census, but 20 years or more of residence could be assumed for migrants present in the 1971 Census. The higher mortality, adjusted for socio-economic position, of migrant South Asians aged 45–64 in 1991 who arrived before 1971 compared with that of those who arrived after 1971 illustrated the possible wearing off of the 'healthy worker effect' and also the impact of environmental influences.

9.4 Conclusion

Meaningful ethnic classifications, return migration, the 'healthy migrant effect', the type of migration – whether voluntary or involuntary – and generational differences in mortality, all pose significant methodological challenges to the use of the data for projecting ethnic populations in the UK. The scarcity of relevant longitudinal data, and hence the very few empirical studies from which to derive accurate estimates of the effect of these factors, further compounds the problems. Some of these issues would be addressed if the ethnic populations in the LS were boosted and ethnic group was collected at death registration.

References

Abbotts J, Williams R and Davey Smith G (1998) Mortality in men of Irish heritage in West Scotland. *Public Health*, vol 112 no 4, pp 229–32.

Anwar M (1979) *The myth of return: Pakistanis in Britain*, Heinemann: London.

Arai L and Harding S (2002) UK-born Black Caribbeans: generational changes in health and well-being. a literature review.

Balarajan R, Bulusu L, Adelstein A M and Shukla V (1984) Patterns of mortality among migrants to England and Wales from the Indian subcontinent. *British Medical Journal*, vol 289, pp 1185–7.

Balarajan R and Raleigh V S (1997) Patterns of mortality among Bangladeshis in England and Wales. *Ethnicity and Health*, vol 2 nos 1–2, pp 5–12.

Bhopal R, Unwin N, White M, Yallop J et al (1999) Heterogeneity of coronary heart disease risk factors in Indian, Pakistani, Bangladeshi, and European origin populations: cross sectional study. *British Medical Journal*, vol 319, pp 215–20.

Bundey S, Alam H, Kaur A, Mir S and Lancashire R (1991) Why do UK-born Pakistani babies have high perinatal and neonatal mortality rates? *Paediatric and Perinatal Epidemiology*, vol 5 no 1, pp 101–14.

Byron M (1994) *Post-war Caribbean migration to Britain: the unfinished cycle*, Avebury: Aldershot.

Byron M and Condon S. (1996) A comparative study of Caribbean return migration from Britain and France: towards a context-dependent explanation, *Transactions of the Institute of British Geographers*, vol. 21, no. 1, pp 91–104.

Cabral H, Fried L E, Levenson S, Amaro H and Zuckerman B (1990) Foreign-born and US-born Black women: differences in health behaviors and birth outcomes. *American Journal of Public Health*, vol 80 no 1, pp 70–2.

Chamberlain, M (1997) *Narratives of exile and return*, Macmillan Press: London.

Davey Smith G, Chaturvedi N, Harding S, Nazroo J and Williams R (2000) Ethnic inequalities in health: a review of UK epidemiological evidence. *Critical Public Health*, vol 10 no 4, pp 375–408.

Fang J, Madhavan S and Alderman M H (1996) The association between birthplace and mortality from cardiovascular causes among Black and White residents of New York City. *New England Journal of Medicine*, vol 335 no 21, pp 1545–51.

Grasmuck S and Pessar P (1991) *Between two islands: Dominican international migration*, University of California Press: Berkeley.

Haenszel H and Kurihara M (1968) Studies of Japanese migrants, I: mortality from cancer and other diseases among Japanese in the United States. *Journal of the National Cancer Institute*, vol 40 no 1, pp 43–68.

Harding S (2000) Examining the contribution of social class to high cardiovascular mortality among Indian, Pakistani and Bangladeshi male migrants living in England and Wales. *Health Statistics Quarterly 5*, pp 26–8.

Harding S (2002, submitted) Mortality from the Indian subcontinent and duration of residence in England and Wales: a longitudinal study.

Harding S and Balarajan R (1996) Patterns of mortality in second generation Irish living in England and Wales: longitudinal study. *British Medical Journal*, vol 312, pp 1389–92.

Harding S and Balarajan R (2000) Limiting long-term illness among Black Caribbeans, Black Africans, Indians, Pakistanis, Bangladeshis and Chinese born in the UK. *Ethnicity and Health*, vol 5 no 1, pp 41–6.

Harding S and Balarajan R (2001) Mortality of third generation Irish people living in England and Wales: longitudinal study *British Medical Journal*, vol 322, pp 466–7.

Harding S, Dews H and Simpson S L (1999) The potential to identify South Asians using a computerised algorithm to classify names. *Population Trends* 97, pp 46–9.

Harding S and Maxwell R (1997) Differences of mortality of migrants, in Drever F and Whitehead M (eds), *Health inequalities*, OPCS series DS no15, TSO: London.

Hattersley L and Creeser R (1995) *Longitudinal Study 1971–1991: history, organisation and quality of data*, OPCS series LS no 7, HMSO: London.

Heath A and Ridge J (1983) Social mobility of ethnic minorities. *Journal of Biosocial Sciences (Supplement)*, vol 8, pp 169–84.

Hummer R A, Biegler M, de Turk P B, Forbes D, Parker Frisbie W, Hong Y and Pullum S G (1999) Race/ethnicity, nativity and infant mortality in the United States. *Social Forces*, vol 77 no 3, pp 1083–18.

Kliewer E (1992) Epidemiology of diseases among migrants. *International migration*, vol 30, pp 141–64.

McMichael A J, Bonett A. (1981) Cancer profiles of British and Southern European migrants. *Medical Journal of Australia*, vol 2, pp 29–232.

Marmot M G, Adelstein A M and Bulusu L (1984) *Immigrant mortality in England and Wales 1970–78: causes of death by country of birth*, OPCS series SMPS no 47, HMSO: London.

National Institute for Ethnic Studies in Health and Social Policy (1997) *Ethnic diversity in England and Wales. an analysis by Health Authorities based on the 1991 Census*, NIESH: London.

Nazroo J Y (1997) *The Health of Britain's ethnic minorities*, Policy Studies Institute: London.

Office of Population Censuses and Surveys and General Register Office for Scotland (1993) *1991 Census: ethnic group and country of birth*, Great Britain, vols 1–2, HMSO: London.

Pallotto E K, Collins J W and David R (2000). Enigma of maternal race and infant birth weight: a population-based study of US-born Black and Caribbean-born Black women. *American Journal of Epidemiology*, vol 151 no 11, pp 1080–5.

Powell J E; Kelly A M, Parkes S E, Cole T R and Mann J R (1995) Cancer and congenital abnormalities in Asian children: a population-based study from the West Midlands. *British Journal of Cancer*, vol 72 no 6, pp 1563–9.

Powell J E, Parkes S E, Cameron A H and Mann J R (1994) Is the risk of cancer increased in Asians living in the UK? *Archives of Disease in Childhood*, vol 71 no 5, pp 398–403.

Razum O, Zeeb H, Akgun H S and Yilmaz S (1998) Low overall mortality of Turkish residents in Germany persists and extends into second generation: merely a healthy migrant effect? *Tropical Medicine and International Health*, vol 3 no 4, pp 297–303.

Razum O, Zeeb H and Rohrmann S (2000) The 'healthy migrant effect' – not merely a fallacy of inaccurate denominator figures. *International Journal of Epidemiology*; vol 29 no 1, pp 191–2.

Sheth T, Nair C, Nargundkar M, Anand S and Yusuf S (1999) Cardiovascular and cancer mortality among Canadians of European, South Asian and Chinese origin from 1979 to 1993: an analysis of 1.2 million deaths. *Canadian Medical Association Journal*, vol 161 no 2, pp 132–8.

Stuart A (1990) *The measurement of ethnicity*, LS User Guide no 2, Centre for Longitudinal Studies, Institute of Education: London.

Weitoft G R, Gullberg A, Hjern A and Rosen M (1999) Mortality statistics in immigrant research: method for adjusting underestimation of mortality. *International Journal of Epidemiology*; vol 28 no 4, pp 756–63.

Wild S and McKeigue P (1997) Cross sectional analysis of mortality by country of birth in England and Wales. *British Medical Journal*, vol 314, pp 705–10.