

Population Estimates by Ethnic Group: Methodology Paper

Summary

This paper describes the methodology used to produce estimates of the population by ethnic group for local authorities in England and Wales. The approach used is a cohort component methodology with population counts, and each component of population change, constrained to the Office for National Statistics (ONS) national and local authority mid-year population estimates. Consideration is given to the modelling of the ethnic dimension of mortality; fertility (and the allocation of ethnic group to infants); switching between ethnic group categories; and the various aspects of migration, with particular attention given to the application of commissioned Census data. The methodology described is that used for Release 8 of the estimates published in May 2011. The methodology has changed very little since the publication of the first set of estimates in 2006.

The paper consists of two parts. The first provides a description of the methodology. The second provides guidance on using the estimates in an informed way.

We welcome any comments on the methodology, estimates or the supporting information provided.

**Population Estimates by Ethnic Group Team
May 2011**

Contact

Email: epe@ons.gov.uk

Tel: 01329 444690

Part 1 Methodology

Introduction

There is increasing and substantial interest in up-to-date estimates of the population of ethnic groups. Some previous estimates have been based on Labour Force Survey results and are thus restricted to high levels of aggregation of geography or ethnic group. Detailed results for each ethnic group, by sex and quinary age-group, were produced as standard output from the 2001 Census in May 2003, but necessarily fail to reflect rapid growth in some groups since 2001.

This article describes a methodology for producing estimates by ethnic group using an orthodox cohort component methodology. The current published estimates cover the period mid-2001 to mid-2009 and are consistent (in both numbers and, where possible, methodology) with the national and local authority mid-year population estimates published by ONS. The methodology constructs estimates for single year of age, sex, and ethnic group at local authority level, though published estimates are aggregated across at least one of these dimensions.

Cohort Component model

The standard approach to producing population estimates is the cohort-component method. This is the method used for the ONS national and local authority mid-year estimates and described in *Making a population estimate*¹. The overall approach is summarised in p11 of that document as follows:

"Summary of the cohort component method

*Take the previous mid-year resident population and age-on by one year;
Then estimate the population change between 1 July and 30 June by;
Adding births occurring during the year
Removing deaths occurring during the year;
Allowing for migration to and from the area*

In addition to the process summarised above, adjustments are also made for some special population groups that are not captured by the internal or international migration estimates: members of the armed forces, prisoners and pupils in boarding schools. These populations have specific age structures, which remain fairly constant over time. Therefore these groups are not aged-on with the rest of the population. "

The cohort component method as applied to population estimates by ethnic group has the advantages that it:

- is consistent with the mid-year estimate methodology;
- allows estimates for small groups to be produced;
- can be extended easily to produce projections consistent with the ONS subnational population projections;
- allows analysis of the relative importance of the components of population change for each ethnic group.

The adoption of the cohort component approach requires the development of a variety of demographic rates and propensities specific to each ethnic group. The methods used to derive these factors are discussed below. The approach places great reliance on using the results of the 2001 Census to identify differences between ethnic groups, and Appendix 1 provides a list of commissioned Census tables used for this work.

Definition of Ethnic Group

The complexities of defining and describing ethnic group are discussed in *Ethnic group statistics*³. For the purposes of this article please note that:

- Ethnic group is self-assigned - that is, chosen by the respondent from a list of categories (including an 'other' option)
- The classification used in National Statistics is the 16-way classification adopted in the 2001 Census (see, for example, Table 1)
- A person's ethnic group can change over time
- Description of ethnic group can change in different contexts. Reliance on the Census data in the modelling process has the *de facto* effect that the estimates will accord with the context of the Census – in particular, this will reflect any effect due to proxy responses by the form-filler on behalf of another household member

Mid-2001 Base Population for England

Before discussing components of change it is necessary to estimate the starting population for the estimates, known as the base population. Following the approach of the mid-year estimates, the initial base population for the estimates is taken to be the 2001 Census population. A specially commissioned table provides Census counts by ethnic group, sex and single year of age for each local authority in England¹. Table 1 provides a summary of the Census results for each ethnic group.

¹ Data taken from Table C0533. Separate counts for the City of London and Isles of Scilly are estimated using C0533 in conjunction with Table S102.

Table 1: Population by ethnic group: England, 2001 Census

	Total (thousands)	% of total	Median age*	% of population female and aged 15-44
All people	49,139	100.0	37	40.7
White				
White: British	42,747	87.0	39	39.1
White: Irish	624	1.3	51	33.5
White: Other White	1,308	2.7	33	57.3
Mixed				
White and Black Caribbean	231	0.5	13	40.1
White and Black African	76	0.2	18	46.9
White and Asian	184	0.4	17	43.3
Other Mixed	151	0.3	18	46.3
Asian or Asian British				
Indian	1,029	2.1	31	52.5
Pakistani	707	1.4	22	51.5
Bangladeshi	275	0.6	21	50.8
Other Asian	238	0.5	30	51.4
Black or Black British				
Caribbean	561	1.1	35	52.7
African	476	1.0	27	59.2
Other Black	95	0.2	22	55.9
Chinese or other ethnic group				
Chinese	221	0.4	29	59.1
Other Ethnic Group	215	0.4	31	61.8

Source: 2001 Census, Tables KS06, C0533

* Please note that the calculated figures shown here, which are consistent with published Census data, differ to the figures shown in previous methodology notes which were calculated differently

The base population is rolled forward to mid-2001 using the same methodology used for later years as described below. In addition, further adjustments are made to the mid-2001 counts to correspond with adjustments made in the mid-year estimates.

Mid-2001 Base Population for Wales

As the population estimates by ethnic group for Wales were produced at a much later date than the equivalent estimates for England, a number of the data files used for identifying the components of change from the 2001 Census to mid-2001 for England were not readily available for Wales. A pragmatic approach was taken to use proportions of the population by age and sex by ethnic group for Welsh local authorities from the 2001 Census and to apply these to the mid-2001 local authority population estimates, thereby producing mid-2001 population estimates by ethnic group.

Table 2: Population by ethnic group: Wales, 2001 Census

	Total (thousands)	% of total	Median age	% of population female and aged 15-44
All people	2,903.0	100.0	39	38.2
White				
White: British	2,786.6	96.0	39	37.8
White: Irish	17.6	0.6	52	31.9
White: Other White	37.2	1.3	38	48.3
Mixed				
White and Black Caribbean	6.0	0.2	16	41.0
White and Black African	2.4	0.1	18	42.6
White and Asian	5.0	0.2	16	44.4
Other Mixed	4.3	0.1	20	43.0
Asian or Asian British				
Indian	8.3	0.3	30	58.0
Pakistani	8.3	0.3	24	49.3
Bangladeshi	5.4	0.2	20	51.4
Other Asian	3.5	0.1	32	50.4
Black or Black British				
Caribbean	2.6	0.1	39	51.1
African	3.7	0.1	27	51.5
Other Black	0.7	0.0	30	57.0
Chinese or other ethnic group				
Chinese	6.2	0.2	29	58.2
Other Ethnic Group	5.1	0.2	31	66.6

Source: 2001 Census, Tables KS06, C0996

Subtraction of Special Population

As noted above, certain special population groups are subtracted from the mid-year estimates before ageing-on the population and applying the components of change. A similar approach is adopted in the population estimates by ethnic group.

Armed Forces

In general, the ethnic composition of the Armed Forces in an area is estimated by applying the ethnic composition of Armed Forces in that area recorded in the Census to the total Armed Forces population used in the mid-year population estimates. Home and Foreign Armed Forces (the latter includes dependants) are treated separately in recognition of the different ethnic profiles of the two populations. From mid-2005 onwards the dependants of Foreign Armed Forces are no longer considered a special population due to changes in the way they move in and out of the country. These flows in and out of the country are now captured through the International Passenger Survey (IPS). The dependants, as at the end of mid-2005, have become part of the standard population.

Prisoners

A similar approach is taken in estimating the ethnic composition of prisoners. Again, the ethnic composition of prisoners in that area recorded in the Census is applied to the population of prisoners used in the mid-year population estimates. Calculations are carried out separately for male and female populations.

School Boarders

As no reliable information on the ethnic group of school boarders is available, it is assumed that these share the ethnic characteristics of people of that age and sex in that area.

Components of change and ethnic group

This section describes the methods adopted in estimating the various components of change in the model. Detailed discussions of issues relating to each component of change are contained in *Population projections by ethnic group: A feasibility study*².

Mortality

The standard method of calculating mortality rates is to use counts of death from the death register and estimates of the population at risk from the mid-year estimates. However, neither of these sources includes data on ethnic group. Studies using country of birth as a proxy for ethnic group are becoming less informative as in-migrant populations move to second or third generation. Analysis of ONS Longitudinal Study data did not provide evidence on which differences in mortality rates between groups could be reliably assumed.

The methodology adopted thus takes the age-specific mortality rates estimated for each area using registered deaths and the estimated mid-year population and applies them to each ethnic group. It will be noted that, as these rates vary by area, and ethnic groups are not distributed evenly across areas, this method will produce implied different mortality rate profiles for each ethnic group across England and Wales as a whole.

Fertility

Age-specific fertility rates (ASFRs) will be an important determinant of relative growth rates of ethnic groups. It seems intuitively likely that cultural factors, (and the immigration of spouses), have a substantial effect on both the number and timing of births.

The conventional method of estimating fertility rates is to divide the number of births by women of a particular age (provided by birth register statistics) by the number of women of that age in the population (derived from the mid-year estimates). This approach is not immediately possible for estimates by ethnic group as the birth register does not record ethnic group of mother (note: fertility is here discussed in the context of the ethnic group of the mother: ethnic group of the child is considered below) and the mid-year estimates do not provide separate counts by ethnic group to act as the denominator for the fertility rate.

The method adopted uses 2001 Census data on the age and ethnic group of mothers of 0-year-olds, together with the counts of all women of that age in that ethnic group, to derive an estimated 'mothering ratio' for each ethnic group. Conceptually, these ratios are divided by the overall ratio to provide the differential profile for each group. Since a mother (of an infant) aged x at Census day may have been aged x or $x-1$ at the time of the infant's birth, the differential fertility profile for each group is approximated as the mean of the mothering ratios for the two applicable years. These profiles are estimated for ages 15 to 44, with the profile for each group then scaled up to allow for the small proportion of mothers aged outside this age band. To allow for the possibility of different patterns in differential fertility

within England and Wales, the calculation of differential fertility profiles is carried out separately for London and the rest of England and Wales.

Example: Calculation of the number of births to women within an ethnic group

In local authority A (in London) there are 200 Asian Bangladeshi females aged 23 and 1,000 White British females aged 23; and no other females aged 23 (at end of year).

The (Census based) estimated fertility rate for Asian Bangladeshi (aged 23) in London is 0.05. The (Census based) estimated fertility rate for all ethnic groups (aged 23) in London is 0.025.

Thus 23 year old Bangladeshi women in Local authorities in London boroughs will initially be assumed to have an ASFR twice as high as the average for that area.

The (mid-year estimate based) estimated age-specific fertility rate for 23 year old women in local authority A is 0.06

Therefore, the initial estimate of the ASFR for 23 year old Bangladeshi women is $0.06 \times 2 = 0.12$

And the number of births to Bangladeshi women aged 23 is $200 \times 0.12 = 24$.

Carrying out similar calculations for all ages and ethnic groups produces an estimate of 500 births in local authority A.

The birth registration figures show that the actual number of births in the area was 400. Thus, the scaling factor = $400/500 = 0.8$.

The scaled estimate of births to Asian Bangladeshi women aged 23 in local authority A is thus $0.8 \times 24 = 19.2$

It should be stressed that the fertility rates are initially applied to the female population before taking account of mortality and migration. This approach is consistent with the ONS subnational population projections but has the weakness that slightly too few births will be generated for groups which have relatively high net in-flows of women of child-bearing age (and similarly slightly too many births generated for groups with relatively low net in-flows). In addition, the 7.1 per cent of infants in households (and infants in Communal Establishments) not linked with their mother on the Census records will not be reflected in the estimates of fertility differentials. This latter weakness would lead to an underestimate of fertility rates for groups with a disproportionate number of mothers not linked with their infants.

This weakness has been avoided in the revised methodology used for the August 2006 and later releases, where fertility rates are calculated to take account of infants not linked to their mother. In practice, these weaknesses are likely to have a relatively small impact on the quality of the estimates compared to the uncertainty in the estimates of international migration.

The above calculation allows estimates to be made of the number of babies born to women of each ethnic group. However, to estimate the number of babies of each ethnic group account must be taken of heteroethnic infancies – that is, the propensity for mothers to have different ethnic characteristics from their children. This is done using factors derived from

Census data linking ethnic group of mother to ethnic group of child - for example, showing that 5 per cent of White: Other 0-year-olds have a White British mother.

These factors can be expected to change over time as the ethnic composition of the population changes. This effect is reflected, to some extent, by calculating similar factors for children aged 1, 2, 3, and 4, and assuming that it is appropriate to linearly extrapolate factors for 2002 and subsequent years.

Calculating such factors is problematic for individual local authority districts, where many combinations of mother's and infant's ethnic group will contain very small numbers. Thus, it was initially assumed that modelling separate rates for London and for the rest of England and Wales provided an appropriate level of detail. However, this approach would fail to take into account the effects of different concentrations of ethnic groups as potential fathers in different areas within London or the rest of England and Wales. This has a small impact on the estimates for some groups in some areas, and small adjustment is made by applying further factors to the estimated number of births in each ethnic group in each local authority.

These factors are derived by comparing the estimated distribution of births using the above methodology for that part of 2001 between Census day and mid-year with the Census distribution. While this is not exact (the timing is different and the Census includes migrants, for example) it should provide factors which adjust the modelled data closer to reality. The factors are further adjusted to ensure that their application does not change the overall number of births in each ethnic group for England and Wales as a whole. Further research on this aspect of the estimates is planned.

Ethnic Switching

An interesting aspect of modelling population by ethnic groups is the possibility of changes in ethnic affiliation. Some research into this was conducted using Longitudinal Study data for people included in both 1991 and 2001 Censuses. Unfortunately, the difference in ethnic classifications used in the two censuses (with, for example, no 'mixed' categories included in the 1991 classification) makes it difficult to identify genuine changes of affiliation over time. A more detailed investigation of stability in ethnic group affiliation has been provided by Platt, Simpson and Akinwale⁴. Although the model has been set up to allow for the incorporation of such a switching effect, it is assumed that change in ethnic affiliation is not a significant effect in demographic changes in ethnic groups.

Domestic Migration

The estimation of migration between areas within England and Wales is the most involved part of the methodology. It can be broken down into four steps:

First, the numbers of migrants of each ethnic group from each local authority are estimated by applying an age-migration propensity profile for that group to the current population. These profiles are estimated from Census data showing the number of people of each ethnic group and quinary age who moved between local authorities in England and Wales.

Second, these notional migrants are allocated to a destination local authority using Census data on the origin-destination patterns of people of each (quinary) age.

Third, these flows are adjusted to allow for higher/lower flows of some ethnic groups to particular destinations (the so-called 'ethnic effects'). The ethnic effects are estimated by comparing Census data on flows of each ethnic group into each local authority with the flow that would be expected based solely on the age and geographical distribution of that ethnic group

Finally, the matrix of flows by single year of age and sex (summed by ethnic group) is constrained to the matrix of flows for that year used in the mid-year Estimates.

Migration to and from other parts of the UK

Inflow

The ethnic composition of in-migration from each other part of the UK is assumed to be the same as that for the relevant quinary age group in the 2001 Census. Results are calculated separately for each part of the UK (so a large increase in-migration from Northern Ireland, say, would be accompanied by a corresponding increase in the number of White: Irish in-migrants).

In contrast to the international and internal migration components, no attempt is made to reflect differential propensities to migrate to different local authorities by ethnic group. The relatively ethnically homogenous nature of the populations of the other parts of the UK means that attempts to model such effects would be based on very small counts (at the local authority level) and would be unlikely to materially change the estimates.

Outflow

Calculation of ethnic differentials in propensities to migrate to other parts of the UK is complicated by the use of different ethnic classifications in the Scottish and Northern Ireland censuses, and the important influence of location in determining migration destination. The method adopted is as follows:

For Scotland, an arbitrary age-migration curve is applied to the population of each area, with the estimated number of out-migrants by age and sex then constrained to the mid-year estimate figure for the area. Thus it is assumed that there is no difference in the probabilities of migrating to Scotland (say) for the various ethnic groups (for a given sex, age, and local authority of residence).

For Northern Ireland a different approach is adopted to allow for the expected greater probability of White: Irish to migrate to the area. In essence, it is assumed that the ethnic distribution of out-migrants to Northern Ireland is similar to that of in-migrants. The estimated number of out-migrants in each ethnic group (across England and Wales as a whole) in 2001 is divided by the Census population by ethnic group to estimate the proportion of each group which would migrate to Northern Ireland. These proportions are used to scale the standard age-migration curve.

As with Scottish migration, the results of applying the standard curve are then constrained to the mid-year estimate local authority/age/sex totals for migration to Northern Ireland.

The application of propensity to migrate curves means that changes in the ethnic composition of out-migration are a natural result of changes in the ethnic composition of the resident population. While assumptions made on cross-border flows are unlikely to greatly affect the estimates the additional assumptions made to model the White: Irish component of out-migration to Northern Ireland protect against systematic underestimation of this flow which would result from applying the methodology used for Scotland.

International migration

International migration is treated as having four components: migration measured by the International Passenger Survey (IPS), Visitor Switchers, Asylum Seekers and migration to and from Ireland (formerly referred to as the Republic of Ireland). Each of these components has both an inflow and an outflow. The methodology for each component is set out below.

a) IPS/Visitor Switcher migration

The IPS provides a measure of the number of people migrating to England and Wales who stay for at least 12 months. For the purposes of this article 'IPS migration' will be taken to refer to this flow after correction for temporary visitors who stay longer than initially planned (so called 'visitor switchers') and intended migrants who leave before 12 months ('migrant switchers'). In the absence of any evidence that the assumption is unjustified, the ethnic composition of visitor switchers is assumed to be the same as that of intended migrants measured by the IPS.

Modelling the ethnic group of this component of international migration relies on the IPS data on country of birth of migrants and Census data on the relationship between country of birth and ethnic group.

Inflow

The 'IPS inflow' by age and sex into each local authority has already been estimated for the mid-year population estimates. The ethnic composition of these flows is estimated as follows.

Firstly, IPS data on Country of Birth (COB) of in-migrants is combined with a Census distribution of COB against ethnic group. This provides an estimate of the ethnic composition of the total IPS inflow. Applying these factors directly to the total inflows used in the mid-year estimates for each local authority would fail to reflect differential propensities of ethnic groups to migrate into a particular district (for example, the Census suggests that Bradford attracts 7 per cent of all international in-migrants of the Asian: Pakistani ethnic group). This effect is dealt with as follows.

The proportion of international migrants of each ethnic group going to each local authority is calculated using Census data. These proportions are applied to the England and Wales level estimates by ethnic group described above to produce initial estimates of the ethnic group international in-migrants in each local authority. These estimates (by age, sex, ethnic group and local authority) are then scaled back to the mid-year estimates flows for that age, sex and local authority. The methodology was revised for the August 2006 release to account for particular migration effects in areas with a large number of US or Gurkha Armed Forces.

The adjustment for differential propensities of ethnic groups to migrate into a particular district results in final estimates of IPS in-migration by ethnic group which do not accord precisely with the initial estimates. This discrepancy is removed using iterative proportional fitting to allow consistency with both the IPS-derived estimates of ethnicity and the IPS in-migration constraints used in the mid-year Estimates.

Outflow

The calculation of the ethnic composition of 'IPS outflow' is simpler than that for inflow. Again, information on country of birth of migrants is used to estimate the ethnic composition of the outflow from England and Wales as a whole. These estimates are divided by estimates of the size of the total population of that group to produce a measure which can be most easily understood as a probability of a person of that group emigrating. These 'probabilities' can thus be applied to the populations of each group within each local authority to provide initial estimates of the number of people of each age, sex and ethnic group within each local authority who emigrate. As with other components, these counts, summed across ethnic groups, are then scaled to the counts used in the mid-year estimates. The method of calculating the out-migration probabilities were improved for the August 2006 and later releases.

A natural criticism of this method is that the relationship between country of birth and ethnic group is unlikely to be the same for emigrants as for all residents. For example, it would be expected that, of those people of a given age born in the UK, those of the Asian Pakistani ethnic group would be more likely to travel to Pakistan than those of the White British group. Although this criticism is accepted, there are two mitigating factors which should be considered. Firstly, the COB-ethnic group mapping used is that used in the calculation of inflow. Where emigration is not permanent, then, an underestimate of non-‘White British’ group emigrating to a particular country should be mirrored by an underestimate of that group returning from that country (although the two flows would occur at different times). Secondly, the flows of UK-born persons to the Caribbean Commonwealth and the Indian sub-continent, which are those generally identified as being of concern in this context, are relatively small as set out in Table 3.

Table 3: IPS emigration of people born in UK (selected next country of residence): UK, 2003

Country of next residence	Outflow (thousands)
All	162.3
European Union	67.0
Australia	36.7
New Zealand	10.4
Bangladeshi, India, Sri Lanka	0.4
Pakistan	1.9
Caribbean Commonwealth	0.0

Source: Table 3.20, MN30 International Migration, ONS

b) Asylum Seekers

The estimation of the ethnic composition of asylum-seeker flows is based on combining the detailed nationality figures for net flows of asylum-seekers (including both Principal Applicants and Dependants) for each year with the Census cross-tabulation of country of birth and ethnic group. This ethnic composition is then applied to the flows, by age and sex, into, or from, each area. For the October 2007 release, this approach was refined to model inflows and outflows separately. As a matter of practicality, and in contrast to the calculation of the characteristics of IPS migration, it is assumed that no asylum-seekers are White British. Whilst this assumption can scarcely fail to be incorrect, it is likely to reflect the composition of asylum-seeker flows more accurately than the raw country of birth - ethnic group data (which typically show substantial proportions of people with each country of birth having White: British ethnicity).

Several points should be made in reference to the estimation of this component.

Firstly, the assumption that the 2001 Census data on country of birth is a fair proxy for nationality of asylum seeker should be acknowledged. This assumption can be criticised on several grounds - country of birth is, of course, different from nationality (and this is a prime reason why the *ad hoc* assumption that no asylum seekers are White: British is made); it does not take account of changes in the ethnic composition of a country between the initial migration (of the population with that country of birth recorded in the 2001 Census) to England and Wales and the asylum-seeker flow; and it does not allow for the possibility that ethnic group is itself a prime determinant of whether somebody of a particular nationality becomes an asylum-seeker.

Secondly, whilst the distribution of asylum-seekers between local authorities follows that used in the mid-year estimates, the additional assumption is made that the calculated ethnic distribution of asylum seekers for England and Wales applies for each local authority (thus, if

10 per cent of asylum seekers were Asian: Pakistani, for example, 10 per cent of the asylum seeker flow into each local authority will be Asian: Pakistani).

c) Irish Flows

These small flows are disaggregated by assuming an ethnic composition for both inflows and outflows similar to that of inflows from Ireland recorded in the Census. From 2009, flows to and from Ireland are reflected in IPS data and are thus subsumed in the IPS migration component described above.

Further Adjustments to Mid-2001 Counts

Although the estimates by ethnic group are based on the 2001 Census, they do incorporate the revisions made to the mid-year estimates made since the Census results were first published. These adjustments are described below.

Unprocessed forms

Corrections were made in the mid-year estimates for about 5,100 people in England and Wales who were included on unprocessed Census forms. As these forms were largely concentrated in particular wards within certain local authorities, the ethnic composition of this adjustment is assumed to be the same as the ethnic composition for that age/sex group within that ward.

Longitudinal Study adjustment

This major adjustment, of about 164,000 sought to correct a believed underestimation of (mostly) males aged 25 to 34. It is assumed that the ethnic composition of this adjustment is the same as the ethnic composition for that age/sex group within that local authority. The possibility of further research on this component using Longitudinal Study data is being considered.

Local Authority Studies (including Manchester adjustment)

Adjustments to the estimates for 15 local authorities in England and Wales were made following the detailed Local Authority Studies. These adjustments totalled 107,000. As with the Longitudinal Study adjustment, it is assumed that the ethnic composition of this adjustment is the same as the ethnic composition for that age/sex group within that local authority.

More information on these adjustments in the mid-year estimates is available at <http://www.statistics.gov.uk/about/data/methodology/specific/population/PEMethodology/>.

Population Estimates by Ethnic Group for Primary Care Organisations (PCOs)

In September 2008 population estimates by ethnic group statistics were published at PCO level for the first time (covering England only). Most PCOs are formed from one or more whole local authorities but some (21 out of 152 PCOs) are formed, wholly or in part, from split local authorities. 2001 Census data was used to calculate the proportion of each ethnic group (by age and sex) in each part of the split local authority. These proportions were applied to local authority level estimates and the resulting data are aggregated at the PCO level. The final estimates are calculated by constraining to the relevant mid-year estimates by PCO.

Reliability and Variability

The reliability of estimates produced using the above methodology is difficult to quantify owing to the nature of potential sources of error.

Firstly, there is uncertainty inherent in the mid-year population estimates, to which the estimates by ethnic group are constrained. This uncertainty encompasses, *inter alia*, various sources of variability in the 2001 Census counts³; limitations in estimating internal migration from administrative records; and the effect of basing estimates of international migration on sample surveys.

Secondly, assumptions on appropriate proxies may be incorrect. In particular, the attribution of ethnic group to international migration flows is predicated on the assumption that the country of birth-ethnic group distribution recorded in the 2001 Census for existing residents can be appropriately applied to flows of people with that country of birth or, with asylum-seekers, the associated nationality. Further detail on the assumptions underlying the estimates is provided in the *Issues and Guidance* paper, available at <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14238>

Thirdly, reliance on the 2001 Census data to identify differences in demographic rates between ethnic groups can be expected to become less adequate through the inter-censal period.

The robustness of the estimates to errors in estimated parameters is summarised in Table 4, which shows the effect on the estimate of the total population of an ethnic group of a 1 per cent error in the initial estimated flow for that group (that is, before constraining to the mid-year Estimate total for that component). These alternative scenarios are run for 2002-2003 using the published 2002 results, and are compared with the published 2003 estimates (this analysis is based on the estimates published in 2006 but would change very little if repeated with estimates from Release 8 published in May 2011). The table shows, for example, that if assumed mortality rates were increased by 1 per cent at all ages for the Asian: Indian group, holding all other rates constant, the estimate for that group would be 0.004 per cent lower. Small and opposite effects would be seen in other ethnic groups, where the number of deaths would reduce to ensure the total number of deaths remains constant.

Table 4: Sensitivity analysis for 1 percentage error in estimated flows: England, 2003

	Births	Deaths	Flow from rest of UK	Flow to rest of UK	IPS in-migration	IPS out-migration	Asylum Seekers (net flow)	Irish inflow	Irish outflow
White									
White: British	0.002	-0.002	0.000	0.000	0.000	-0.001	0.000	0.000	0.000
White: Irish	0.004	-0.014	0.001	-0.004	0.003	-0.003	0.000	0.004	-0.009
White: Other White	0.009	-0.006	0.001	-0.003	0.062	-0.040	0.001	0.000	0.000
Mixed									
White and Black Caribbean	0.035	-0.002	0.000	-0.002	0.004	-0.003	0.000	0.000	0.000
White and Black African	0.039	-0.002	0.001	-0.003	0.022	-0.010	0.017	0.000	0.000
White and Asian	0.039	-0.002	0.001	-0.003	0.017	-0.008	0.007	0.000	0.000
Other Mixed	0.036	-0.003	0.001	-0.003	0.022	-0.012	0.008	0.000	0.000
Asian or Asian British									
Indian	0.012	-0.004	0.001	-0.002	0.030	-0.010	0.004	0.000	0.000
Pakistani	0.021	-0.003	0.001	-0.002	0.014	-0.006	0.005	0.000	0.000
Bangladeshi	0.022	-0.002	0.000	-0.002	0.017	-0.006	0.002	0.000	0.000
Other Asian	0.015	-0.003	0.001	-0.002	0.036	-0.009	0.018	0.000	0.000
Black or Black British									
Caribbean	0.011	-0.006	0.000	-0.002	0.009	-0.006	0.001	0.000	0.000
African	0.018	-0.002	0.001	-0.002	0.034	-0.011	0.044	0.000	0.000
Other Black	0.025	-0.002	0.000	-0.002	0.012	-0.010	0.005	0.000	0.000
Chinese or other ethnic group									
Chinese	0.009	-0.003	0.001	-0.003	0.116	-0.050	0.020	0.000	0.000
Other Ethnic Group	0.009	-0.002	0.001	-0.003	0.071	-0.022	0.022	0.000	0.000

Dissemination of Estimates

Results are provided for mid-years 2001 to 2009 in six standard tables.

- Table EE1: Population estimates by sex and ethnic
- Table EE2: Population estimates by sex, broad (3 way) age, and (16 way) ethnic group
- Table EE3: Population estimates by sex, broad (3 way) age, and broad (5 way) Ethnic Group
- Table EE4: Population estimates by quinary age by sex and ethnic group
- Table EE5: Components of population change by ethnic group
- Table EE6: Population change by ethnic

Following the practice of the mid-year population estimates counts in table are rounded to the nearest 100.

Tables EE4 and EE5, which would contain very small cell counts if produced for local authority districts, are produced for England and Wales only. All other tables are produced for the standard administrative hierarchy of local authority districts, counties, and regions in England and for local authorities in Wales, and for the Primary Care Organisations and Strategic Health Authorities in England as defined in 2009.

All six tables are also available on the National Statistics website through <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14238>. Tables are also made available through the Neighbourhood Statistics website at <http://neighbourhood.statistics.gov.uk/dissemination/>.

Further Developments

The results of the 2011 Census will provide the first opportunity to check the accuracy of the estimates. At that time we will also examine each part of the methodology to assess whether a better approach is possible – for example using new data sources, or alternative assumptions. We will also rebase assumptions on to the results of the 2011 Census and produce estimates for years following 2011 on the changed ethnic group classification used in the 2011 Census. We expect to put forward those new estimates for assessment for National Statistics status.

Further comments from users on the methodology or the estimates themselves are welcomed.

Part 2 Guidance

General

Status as Experimental Statistics

The Population Estimates by Ethnic Group are experimental statistics. This means that they have not yet been shown to meet the quality criteria for National Statistics, but are being published to involve users in the development of the methodology and to help build quality at an early stage. More information on Experimental Statistics and National Statistics is provided in the National Statistics Code of Practice: Protocol on Data Presentation, Dissemination and Pricing available at http://www.statistics.gov.uk/about/national_statistics/cop/protocols_published.asp. We plan to put forward the estimates for assessment as National Statistics following quality assurance using, amongst other information, the 2011 Census results.

Reliance on 2001 Census Data for Parameter Estimation

Whilst the absolute level of demographic flows for each age and sex are taken from the Mid-Year Population Estimates, the estimation of the ethnic composition of flows is generally achieved using 2001 Census data on ethnic differentials in, for example, fertility. The method is robust to changes in the size of the population of an ethnic group - for example, if the White: Other group grew by 10%, the number of births would increase appropriately, rather than being constrained to the level in the 2001 Census, but cannot track changes in differentials in the demographic rates - that is, if the Census data suggests that White: Other females are 10% more fertile at a particular age than average, that differential is assumed to continue over the period of the estimates (an exception to this is in the allocation of infants to ethnic groups, where changes over the 5 years prior to the Census are extrapolated to provide slightly different factors in subsequent years).

The Census data used in the parameter estimation processes relates to all usual residents. It thus includes people employed in the Armed Forces and prisoners, though these groups are excluded from the demographic transitions applied to the rest of the population. In practice, this will have a negligible effect except for some small effects in LADs with large foreign Armed Forces populations - namely 42UC Forest Heath and, to a lesser extent, 12UC East Cambridgeshire and 33UB Breckland.

Constraint to Mid-Year Estimates

As mentioned above, the components of change applied in the Population Estimates by Ethnic Group are constrained (by single year of age, sex and LAD) to those used in the Mid-Year Population Estimates. Any errors in the Mid-Year Estimates will thus be replicated in the estimates by ethnic group. More information on the methodology used in producing the Mid-Year Estimates is available at <http://www.statistics.gov.uk/about/data/methodology/specific/population/PEMethodology/>.

Base Population and Components of Change

Base Population

The base population is the 2001 Census population. The Census Quality Report, available at <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14212>, provides information on variability due to the estimation of people not appearing on Census forms, the imputation of missing answers, and other sources of error.

Mortality

Assumed mortality rates are the same for each ethnic group within a particular sex/age/LAD group (this does not imply the same mortality rates for each group across England as a whole as a particular ethnic group may have concentrations in LADs with high or low mortality rates). As mortality rates are applied to the population before taking into account migration, there will be a slight under-estimate of deaths in groups with relatively high net inflows (though the concentration of migrants in the younger age groups (with lower mortality rates) mitigates this).

Fertility and Allocation of Births

There are several aspects of the births component of change which should be considered.

Firstly, fertility rates are applied to the mean of the starting population and the aged-on population for each single year of age (before allowing for mortality and migration). This follows the practice of the current Sub-National Projections methodology, but will lead to a slight under-estimate of births in groups with relatively high levels of net cross-border or international in-migration (results for sub-national areas may also be affected by the exclusion of births to migrants from elsewhere in England and Wales). Approximate calculations suggest that this effect in 2003 leads to an under-estimate of around 0.2% for the Other: Chinese and Other: Other Ethnic Group groups, with other groups showing smaller differences. These errors are additive over time.

Secondly, it is acknowledged that differences in estimated mother-infant ratios are not an exact equivalent for differences in fertility rates. However, in contrast to the first release of the Population Estimates by Ethnic Group, the methodology used for the August 2006 and later releases includes a correction for infants not linked with their mothers on a Census form.

Thirdly, LAD-specific differences in ethnic group age-fertility profiles are not modelled. Thus, while the methodology takes account of the fact that, for example, 20 year old Asian Bangladeshi women in London are more likely to have a child than average, and that 20 year old women in Barnet are less likely to have a child than average, it does not take attempt to model the fertility of 20 year old Asian Bangladeshi women in Barnet separately.

Fourthly, the allocation of births by ethnic group of mother to ethnic group of infant is based on extrapolation from patterns recorded in the Census. Further, the adjustment of the allocation of births to take account of differences due to the available pool of partners assumes the population as at Census day 2001. As this adjustment has only a small effect on the numbers of births in each group and area, this assumption will have a vanishingly small effect on the final estimates (and would not affect the estimates for England and Wales as a whole).

Internal Migration

The internal migration system takes into account ethnic differentials in the propensity to migrate, the effect of different age/sex structures within LAD populations, differences in migration levels and patterns from each LAD and differences in the propensity of a particular group to migrate to a specific LAD. It does not attempt to reflect differences in the propensity of a particular group to migrate from specific LADs. Thus if the Asian: Pakistani group is estimated to have a higher propensity to migrate than Asian: Indian across England and Wales as a whole, this will also apply for each LAD (though it is quite possible for a higher proportion of Asian: Indian to migrate from an LAD as a result of the age-sex structures of the groups within that LAD). The calculation of ethnic differentials in the propensity to migrate (across England and Wales as a whole) takes into account differences in the age structure of the population of each group.

Cross-Border Migration

This component covers migration between England and the rest of the UK. The following points should be borne in mind when using the data.

Migration to Scotland

In contrast to the internal migration methodology, the propensity to migrate to Scotland or Wales is assumed to be equal for each ethnic group (after allowing for differences in the age and sex structure and area of residence of each group).

Migration to Northern Ireland

An assumption is made that the ethnic composition of migration to Northern Ireland is similar to the estimated composition of the inflow. This assumption is intended to reflect that the proportion of out-migrants of White: Irish ethnic group is expected to be higher than the proportion in the resident population.

Migration from the rest of the UK

The ethnic composition of flows from Scotland and Northern Ireland (separately) are assumed to be the same as those recorded in the 2001 Census. As the ethnic composition of the populations of the other parts of the UK is likely to change only slightly over the period of the estimates there is no *prima facie* reason to expect the ethnic composition of in-migrants from each area to England to vary substantially (though changes in the age distribution as well as the absolute size of the different groups would affect the composition of the inflow).

International Migration

International migration is modelled as three sets of inflows and outflows.

IPS/VS Flows

The largest inflows of international migration are those measured by the International Passenger Survey. As a sample survey, this is subject to sampling error, which has a quantifiable effect on the variability of the estimates of migration for each ethnic group. A measure of this variability, for 2003, is shown in Table 5.

Table 5: Standard errors of estimates of IPS in migration for ethnic groups

Thousands

	2003 Population estimate	IPS inmigration estimate	Standard error of inmigration estimate	Coefficient of Variation
White				
White: British	42777	1568	101.7	0.065
White: Irish	611	18	0.9	0.049
Other White	1445	1032	70.1	0.068
Mixed				
White and Black Caribbean	250	11	0.8	0.074
White and Black African	90	23	1.4	0.062
White and Asian	209	40	1.9	0.048
Other Mixed	171	43	2.5	0.057
Asian or Asian British				
Indian	1116	380	31.0	0.082
Pakistani	770	125	15.5	0.124
Bangladeshi	304	58	12.8	0.223
Other Asian	279	115	12.4	0.108
Black or Black British				
Black Caribbean	581	61	10.5	0.173
Black African	584	228	18.8	0.083
Other Black	104	14	0.9	0.062
Chinese or other ethnic group				
Chinese	288	376	36.8	0.098
Other Ethnic Group	279	223	21.5	0.096

Source: Population Estimates by Ethnic Group (based on estimates released in January 2006)

Note: This table shows standard errors of IPS in-migration estimates for people by country of birth mapped to ethnic group using the Population Estimates by Ethnic Group methodology. Estimates are those used in the January 2006 release but the analysis would change very little if repeated with estimates from Release 8 published in May 2011.

In addition to sampling variability, the estimation of migration by ethnic group relies on the Census mapping of country of birth to ethnicity being an appropriate method of allocating ethnic group to new migrants.

Asylum Seeker Flows

The ethnic composition of flows of asylum seekers is estimated using Home Office data on the nationality of applicants and dependants, and the Census mapping of country of birth to ethnicity, and assumes that country of birth is an appropriate proxy for nationality. It is assumed that no asylum-seekers have White British ethnicity.

A natural criticism of these methods is that the relationship between country of birth and ethnic group is unlikely to be the same for emigrants as for all residents. For example, it would be expected that, of those people of a given age born in the UK, those of the Asian Pakistani ethnic group would be more likely to travel to Pakistan than those of the White British group. Although this criticism is accepted, there are two mitigating factors which should be considered. Firstly, the COB-ethnic group mapping used is that used in the calculation of inflow. Where emigration is not permanent, then, an underestimate of non-White British group emigrating to a particular country should be mirrored by an underestimate of that group returning from that country (although the two flows would occur at different times). Secondly, the flows which are generally identified as being of concern in this context are relatively small as illustrated in Table 6. The further assumption that no asylum seekers have White: British ethnicity can scarcely fail to be incorrect, particularly with reference to flows from countries such as Zimbabwe. However, the assumption is made here that there is no defensible assumption on the likelihood of an asylum seeker having White British ethnicity which would make the estimates more accurate.

Table 6: International emigration of people born in UK (selected next country of residence): UK, 2003

Country of next residence	Outflow (thousands)
All	162.3
European Union	67
Australia	36.7
New Zealand	10.4
Bangladeshi, India, Sri Lanka	0.4
Pakistan	1.9
Caribbean Commonwealth	-

Source: Table 3.20, MN30 International Migration, ONS

There is assumed to be no difference between ethnic groups in their assignment to local authority districts (thus, if LAD A receives 2% of asylum seekers, it would also receive 2% of White: Other asylum seekers, 2% of Black: African asylum seekers and so on).

Irish Flows

As with flows to Northern Ireland it is assumed that flows to Ireland have a similar ethnic composition to flows from Ireland. This assumption is removed from estimates relating to 2009 where flows to and from Ireland are reflected in the IPS data.

Prisoners

The proportions of prisoners in each ethnic group within each area are assumed to remain the same as in the Census. Theoretically, this would be expected to underestimate the numbers of prisoners from relatively fast growing groups in post-Census years, though in practice the effect will be insignificant over the period of the estimates.

School Boarders

The Mid-Year Estimates treat School Boarders as a special population, not subject to the demographic transitions of ageing, fertility, death and migration. In the absence of information on the ethnic group of school boarders, a simpler approach is taken the population estimates by ethnic group, with changes in the Mid Year Estimates due to adjustments for boarders simply reflected by constraining to the Mid Year Estimate. Thus for example, if the Mid Year Estimates incorporate a rise of 50 in the number of 14 year old boarders in an LAD, the numbers of 14 year olds in each ethnic group will be scaled up to ensure the estimates by ethnic group accord with the Mid Year Estimates. This simpler approach has two consequences - firstly it is implicitly assumed that the ethnic composition of boarders is similar to that of all people of that sex and age resident in the LAD, secondly, the population of boarders is subject to the demographic transitions described previously. This could have a very slight effect on the ethnic distribution of births for LADs with a large number of female boarders, though in practice the low Age-Specific Fertility Rate for ages to 18 will make the effect insignificant.

The Effect of Errors in Parameter Estimation

With the exception of the IPS international migration data described above, it is difficult to quantify the uncertainty surrounding the estimated components of change for each ethnic group. Table 4, on page 14, does not attempt to do that, but provides estimates of the effect on the published estimates of errors in estimating demographic flows. The table (which is based again on estimates released in 2006) shows, for example, that if assumed mortality rates for 2002/03 were increased by 1% at all ages for the Asian: Indian group, holding all other rates constant, the estimate for that group at mid-2003 would be 0.004% lower. Small and opposite effects would then be seen in other ethnic groups, where the number of deaths would reduce to ensure the total number of deaths remains constant.

Appendix 1: Commissioned 2001 Census Tables

The tables below were commissioned from ONS Census Outputs Branch for the production of these estimates. As with other Census commissioned tables, they are now available free of charge from Census Customer Services.

Commissioned Census tables

C0006	Age by ethnicity for England and Wales
C0009	Sex and age and whether born in UK by ethnicity (migrants to England from Wales)
C0010	Sex and age and whether born in UK by ethnicity (migrants to England from Scotland)
C0011	Sex and age and whether born in UK by ethnicity (migrants to England from Northern Ireland)
C0431	Age, sex and ethnic group by ethnic group of mother for England
C0527	Sex; ethnicity; and age by migration status for England
C0528	Origin and destination of migrants by ethnicity for England
C0529	Age by ethnicity of migrants from England to Scotland
C0530	Age by ethnicity of migrants from England to Wales
C0531	Sex and GOR of residence by ethnicity of international migrants for England and Wales
C0532	Selected country of birth by ethnicity for England
C0533	Sex and age by ethnicity for England
C0534	For England, ethnicity and age of mother of children in households aged 0
C0535	Ethnicity of migrants to England from Ireland
C0981	Persons and age and whether born in UK by ethnicity (migrants to Wales from Scotland)
C0982	Males and age and whether born in UK by ethnicity (migrants to Wales from Scotland)
C0983	Females and age and whether born in UK by ethnicity (migrants to Wales from Scotland)
C0984	Sex and age and whether born in UK by ethnicity (migrants to Wales from Northern Ireland)
C0991	Age, sex and ethnic group by ethnic group of mother for Wales
C0992	Sex; ethnicity; and age by migration status for Wales
C0993	Origin and destination of migrants by ethnicity for Wales
C0995	Selected country of birth by ethnicity for Wales
C0996	Sex and age by ethnicity for Wales
C0997	For Wales, ethnicity and age of mother of children in households aged 0
C0998	Ethnicity of migrants to Wales from Ireland

References

1. Office for National Statistics. *Making a population estimate in England and Wales*, (2005) available through <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=575>
2. Office for National Statistics. *Ethnic group statistics: A guide for the collection and classification of ethnicity data*. London: The Stationery Office (2003)
3. Office for National Statistics. *Population projections by ethnic group: A feasibility study*, London: The Stationery Office (2002)
4. Platt, Simpson and Akinwale. *Stability and change in ethnic groups in England and Wales*, *Population Trends* 121, London: Palgrave Macmillan (2005)
5. Office for National Statistics. *Census 2001 Quality Report*, London: Palgrave Macmillan (2005)