

ONE NUMBER CENSUS STEERING COMMITTEE**Uncertainty Intervals for National Demographic Estimates**

1. The attached research has resulted in proposed uncertainty intervals that provide a range of plausible values for the rolled-forward national population estimates in April 2001. The national level rolled-forward population estimates are the best estimates of the population that are independent of the 2001 Census. The uncertainty intervals do not provide better national level population estimates. They will be used to check that the results of the Census and associated census coverage survey (CCS) are of the right order of magnitude.
2. The intervals cannot be determined with certainty, especially as they include estimates of non-sampling errors that have not been measured and are largely incapable of measurement. The figures reflect the judgements of a number of experts who have provided their opinions. They therefore represent informed opinion rather than scientific fact.
3. **The Steering Committee are invited to:**
 - a) **to note the methodology and identify any problems**
 - b) **to confirm that they find the results plausible**
 - c) **to agree that the results be used to check whether the Census/CCS estimates are of the right order of magnitude.**

**John Charlton and Roma Chappell
Office for National Statistics
June 1999**

Uncertainty Intervals for the 2001 National Demographic Estimates

John Charlton and Roma Chappell

Introduction

1. Demographic estimates can be used to provide an alternative estimate of the population in a Census year. Although of high quality, these estimates would not be expected to be as accurate as the census in 2001. Nevertheless, they provide an estimate that the census together with the census coverage survey would be expected to be close to. Therefore, demographic estimates will be used to compare to the census counts together with the census coverage survey after the census in 2001. The calculation of the confidence intervals for the 2001 census results has been described elsewhere – ONS(ONC(SC))98/05. However the demographic estimates are also subject to uncertainty, and this needs to be taken into account when comparisons are made. This paper is about sources of uncertainty in the 2001 demographic estimate.
2. The 2001 demographic estimates will be based on the ONS mid-2001 rolled-forward population estimate that will be produced using a component method. Uncertainty has been estimated for each component separately in terms of sampling error and possible bias and increased variance due to non-sampling error. The details of the estimation of the uncertainty intervals are given in the appendices to this paper. The synthesis of these into our best estimate of the plausible ranges for the demographic estimates is summarised here, and the full details of this synthesis are tabulated in appendix H to this paper. This is not the final uncertainty interval for 2001. Although it is not expected that the final intervals will differ markedly from those shown here they may be slightly wider if more evidence for some components becomes available.
3. While the range of values given here includes an element for sampling errors where survey data have been used, it also includes very rough assessments of likely non-sampling error, based on the informed opinion of experts in the various fields covered by the components of population estimates. Whilst these opinions reflect expert judgement they do inevitably involve a large degree of subjectivity. The uncertainty intervals given in the paper cannot, therefore, be considered as the definitive official estimates of error around national population estimates.
4. The research to provide an uncertainty interval covers England and Wales only. The methodology is available for use by the General Register Office (Scotland) and the Northern Ireland Statistics Research Agency if they choose to adopt it.

The components of the 2001 national demographic estimates

5. The national demographic estimates in 2001 will be based on the 1981 Census adjusted for under-coverage and rolled forward allowing for births, deaths and net migration since 1981 (see ONC(ONS(SC))97/04). The 1981 Census is used as the starting point for this exercise since this was shown to provide the best base for rolling the estimates forward (see ONC(ONS(SC))97/15).
6. The components of the rolled forward demographic estimate for 2001 are the 1981 Census base and the components of change since 1981. These components include births, deaths, migration and armed forces resident in England and Wales in 2001. The rolled forward estimates are usually produced for the mid-year. A 2001 demographic estimate will be produced that is adjusted to take account of the births, deaths and estimated migration during the time between mid-year and Census day. Any uncertainty associated with taking account of the timing difference is allowed for in the uncertainty intervals for the individual components.
7. The data sources that are used for the components of the rolled forward estimate are the best data that are available on a nationally consistent basis. The data are of high quality but the purpose of this research is to acknowledge that there is no such thing as perfection and to attempt to quantify any errors that may be associated with components of the demographic estimate. Some errors are cumulative over time, others are not.
8. COMPONENTS
Non-cumulative components - any errors are not compounded.
1981 census, adjusted for under/over-coverage - this determines the base figures for rolling forward
Adjustments for the armed forces - these adjustments are made each year based on the best data available at that time.

Cumulative components - annual figures are used, and errors compound each year.
The annual cumulative components are:
Births
Deaths
International migration – based on the International Passenger Survey
Migration with Scotland/ Ireland/ Northern Ireland - based on registrations with the National Health Service Central Register (NHSCR)
Asylum seekers - data from Home Office
Visitor switchers - data from Home Office. A visitor switcher is a person who says they are entering the country as a visitor and then stays, or leaves the country for a year or more although initially not intending to settle abroad.

Results for individual elements

9. Uncertainty intervals have been estimated for each of the components of the 2001 demographic estimates. Where data for a component are based on a survey then sampling error has been estimated. In addition, for each component, an attempt has been made to estimate possible sources of bias and increased variance due to non-sampling error. This has been done by consulting experts using the technique of a Delphi study. The details of the uncertainty intervals are given in annex A-G. Annex H shows the detail of the synthesis of the uncertainty intervals for individual components into an overall uncertainty interval. This synthesis is summarised here. Annex I gives further details of the Delphi study.
10. The work on uncertainty intervals has not identified a need for improvements to the population estimates. Therefore, the mid-2001 demographic estimate will provide the best alternative population estimate for 2001 that is independent of the 2001 Census. It is unlikely, however, that

the demographic national population estimate will be exactly the same as the Census estimate and the uncertainty interval provides a range of plausible values for the demographic population estimates. The use that will be made of the uncertainty intervals after the 2001 Census, in reaching the One Number Census estimates, is outside the scope of this paper.

Summary of Results

11. The detailed calculations underlying the summarised uncertainty interval shown here are given in Appendix H to this paper. The results are that the best estimate of uncertainty is $-220,000$ to $+287,000$. Thus, using the Government Actuary's Department projections for 2001, the population range would be between 52,598,000 and 53,105,000.

References

ONS(ONC(SC))98/05 Brown, Buckner, Chambers, Diamond and Teague

ONC(ONS(SC))97/04 Charlton and Chappell

ONC(ONS(SC))97/15 Charlton and Chappell

Thatcher AR (1984). The 1981 Census. J R Statist Soc A (1984) 147, part 2: 222-232

Census Division (1985). 1981 Census evaluation programme. Population Trends 40: 21- 27.

Office for National Statistics. Birth Statistics 1997. Series FM1 no.26

Linstone HA, Turnoff M (eds). Introduction. The Delphi method: techniques and applications. Reading MA, Addison-Wesley, 3-12, 1975.

Charlton J, Patrick DL, Matthews G, West PA. Resource priorities in Kent: a Delphi study. J Epidemiology Community Health, 1981; 35: 288-292.

Appendix A:

Estimating the uncertainty interval for the 1981 population census

The 1981 census has been described as one of the best censuses conducted in the UK (Thatcher,1985), with an estimated undercount of 0.45 per cent, 214,000 persons with confidence interval for the undercount 151,000 – 277,000 based on the survey (Census Division 1985). In 1981 people were enumerated where they were found on Census night (population present base) but for the purpose of making population estimates people were transferred back to their usual residence using information supplied on the form for usual address. The 1981 base figures took account of births and deaths between census day and 30th June, and made allowance for: census under-enumeration estimated by the PES; 26,000 missing children identified from birth registrations; and 300,000 usual residents who were outside GB on census night and thus not included on a census form. The cohort comparisons made for the One Number Census showed a closer match for the 1981 adjusted census results than for the censuses of 1971 and 1991 ONC(ONS(SC))97/15).

In order to arrive at an uncertainty estimate for the 1981, alternative 1981 population estimate(s) that are independent of the Census are required. The only such estimate that is available is the one that has been produced using a cohort analysis. The cohort analysis was based on following through those born each year (in our case since 1911), and making adjustments for mortality each year by applying cohort life tables for that year. Further adjustments were made for migration between England and Wales and elsewhere.

We took the adjusted census data and the data from the cohort analyses as two independent realisations from a random distribution, and calculated the variance based on these two points. Cohort estimates were very similar to official population estimates, especially for ages 0-39, where the migration data used was essentially the same. Compared with the 1971 and 1991 censuses, the 1981 census counts were closest to the official population estimates and cohort estimates. Bias was assumed to be zero, since adjustments had been made for all known factors. For the age group 0-39 the difference between the population estimates for 1981 and the cohort figures was 0.30 per cent. Because cohort figures for older people would be based on less reliable migration assumptions based on port statistics* pre-dating the International Passenger Survey, we calculated the variance for the entire population assuming the relative difference for the older people would be the same (by grossing up). Thus the standard deviation for the 0-39 age group of 59,375 became 104,375 for the entire population.

Conclusion: The uncertainty interval for the 1981 Census is estimated to be +/- 204,575

* Port statistics were used as follows: up to '1919 Intercontinental citizen passengers to and from UK ports'; from 1920-1963 'Intercontinental migration of UK and commonwealth citizens for permanent residence'. If the pre-1964 migration figures were unreliable then discrepancies might be expected to show up in the 1981 table/graphs in age groups 40 and above, as was seen in the figures.

Appendix B

Uncertainty in birth and death data

In making the inter-censal population estimates for England and Wales the number of births and deaths are taken into account as the natural change component each year. The data used are registration based and, in order to ensure that the event is captured in the correct mid-year, the data reflect the event when it occurred rather than at the date of registration.

Birth and death data of residents that occur in England and Wales are considered to be virtually complete because of the registration laws, and benefits attached to registration. A small number of abandoned babies that are found each year are not included in the population estimates. Using data from the 'abandoned children register', it is possible that there are on average 5 babies a year that are abandoned children. This number is negligible in comparison to the number of births in England and Wales each year. Death data are complete, apart from a very small number of homicides that may or may not be identified years after the event. A body cannot be buried/ cremated unless the death has been registered.

The registration of births and deaths in England and Wales is of events that happen here regardless of the usual residence of the person for whom the event is being recorded. Births to residents of England and Wales that are registered elsewhere are thus excluded, while births registered in England and Wales to mothers whose usual residence is elsewhere are included. The net error arising from this factor is estimated to be negligible. This is because, the number of births to residents of England and Wales occurring outside the country is likely to be of the same order as the number of births occurring in England and Wales to non-residents.

In 1997 there were 373 births in England and Wales to non-residents. There were 9879 births to British nationals that were voluntarily notified to British consulates, British High Commissioners or HM Armed Forces registration centres, but most of these were found to be births to women who are not current residents of the UK. ONS estimates that at any time on average about one hundred thousand women of childbearing age (15-44), usually resident in England and Wales, are temporarily absent overseas. Most of these women are absent for only a short period and it is unlikely that more than a few hundred per year would give birth overseas. Most of the births to residents that occur outside of England and Wales take place elsewhere in the UK. For instance, women are likely to give birth at the nearest hospital to where they live and some Scottish women who live near a large hospital in Berwick-on-Tweed in the north of England have their babies there even though they are resident in Scotland. The number of births during 1997 to residents of England and Wales that were registered in Northern Ireland and Scotland were around 25 and 180 respectively.

As with the birth statistics, the deaths that are based on registration data are of deaths that occur in England and Wales. They include deaths of both residents of and visitors to England and Wales and exclude deaths of residents of England and Wales whilst outside the countries. The deaths that are included in the population estimates include about 1500 deaths a year to non-residents of England and Wales that occur here. Some of these are deaths of residents of Scotland or Northern Ireland. In 1997 there were 400 deaths of residents of Scotland and Northern Ireland that occurred in England and Wales, and slightly fewer, 360, deaths of residents of England and Wales in Scotland and Northern Ireland. If there are fewer residents of England and Wales who die abroad than there are non-residents who die in England and Wales then the number of deaths in the 20 years since 1981 will have been overstated. The possible biases that could be due to place of death being different from place of residence are still being investigated, but are thought to be negligible.

Conclusion: Assume no bias in the total number of births and deaths for now, but use the results from the current study when they are available.

Appendix C: International migration – based on the International Passenger Survey

The International Passenger Survey (IPS) is a continuous, voluntary survey conducted by ONS at principal air and sea routes between the UK and countries outside the British Isles. The IPS does not currently cover routes between the UK and Ireland nor is it likely to cover people coming to the UK seeking asylum or people who officially come for a short visit and then apply for permission to stay. These additional categories are covered by other data sources (see Appendix E and F).

The IPS data are subject to both sampling and non-sampling errors including bias. In the Delphi study that is being reported here, a number of experts were consulted about the possible size of these sources of error. In this report, negative bias indicates that the true value is underestimated, i.e. the estimate equals the true value plus bias.

IPS sampling error can be calculated directly – there is also a design effect that needs to be taken into account. The 20-year sampling standard error (adjusted for design effects) was:

Inflow to England and Wales	59,278 (including design effect on variance = 1.2)
Outflow from England and Wales	48,388 (no design effect)

Non-sampling error is difficult to estimate, even if we carried out the elaborate and expensive studies that would be necessary to estimate these. Our alternative approach was to involve experts in a Delphi study (see Appendix I). This is a technique for obtaining opinions of experts anonymously, and allowing them to review their responses in the light of the views of others.

The experts who participated were John Salt, David Orr, Garnet Compton, Peter Lynn and Ludi Simpson, and from ONS Lucy Vickers, Iain MacLeay, Dave Elliot and Tim Jones.

The results (weighted average for respondents) were: the variance of the estimates could be larger, mainly due to potential interviewer effects. The estimated inflation factors (weighted average of responses) were: inflow+9.5 per cent, outflow +7.5 per cent

The view of some respondents was that there could be biases in the estimates, mainly due to the possibility that individuals might be missed by the survey. The estimated biases were: inflow – 3.0 per cent, outflow –2.7 per cent (negative bias indicates that the true value is underestimated, i.e. the estimate equals the true value plus bias).

For biases high and low variants were also calculated. For the high variant the third largest value was taken as the upper quartile estimate (the two most extreme values were far in excess of other estimates and not representative of most expert's views). This gives a bias figure of –8 per cent for inflow and – 5.5 per cent for outflow. For the low variant the lowest figure was taken, as there was little variability at the lowest end of the range. This gives an estimate of bias of 0.0 per cent for both in-flow and outflow.

Conclusion: The effects of these variant estimates is shown in the summarising table in appendix H. The view of experts was that both flows are likely to be underestimated and the best estimate of the net effect of the bias is a possible underestimate of the population of 25,834. In rounded terms, the net effect of the bias and variability of migration flows is a range of –124,100 to 175,800 (to the nearest hundred).

Appendix D Migration with Scotland/ Ireland/ Northern Ireland - based on registrations with the National Health Service Central Register (NHSCR)

For estimating sources of uncertainty the national experts in the constituent countries of the UK were sent the figures for migration between the UK countries that have been used for England and Wales population estimates, and asked to comment. To date no response has been received from Scotland – the results included in this paper are those of the national experts in England and Wales.

There is no sampling error only non-sampling error since the source of this information is one hundred percent administrative data. Potential non-sampling error was described in terms of a range for bias.

Conclusion: The estimate of uncertainty for flows both to and from England and Wales was +/- 5 per cent, i.e. +/- 61093 on immigration and +/- 64816 on emigration. In producing an estimate of the net effect there are four possible combinations of these uncertainty ranges. The most impact is seen when one flow is underestimated and the other is overestimated. It is not the view of the expert that this is likely to be the case. Rather it is felt that both flows are likely to be under estimated to the same extent giving the best estimate of the possible bias as - 3700.

This uncertainty interval may be updated when the response is collected from GRO(S).

Appendix E: Flows to and from Eire

These flows are relatively small. There has recently been an Eurostat study of these migrations, and one of the participants, Fred Ashwood, produced uncertainty estimates for us based on this study and consultations with the Scottish Office and CSO Dublin.

The credible range estimates were:

Inflow to GB in 1981-91:	270,000-230,000
Outflow GB to Eire 1981-91:	120,000-100,000
Inflow to GB 1991-1998:	125,000-95,000
Outflow GB to Eire 1991-98:	160,000-115,000

We scaled the 1991-98 figures to correspond to 1991-2001 figures by multiplying by 10/7, i.e. assuming that the current flows and uncertainties would continue. For trends beyond 1998 the 1991-98 flows have been used in preference to the flows for the full period from 1981, because sustained growth in the Irish economy led to a change in the pattern of the flows to a net inflow to Ireland in recent years, and it has been assumed that this will continue to 2001. England and Wales figures were taken as 0.96 of the GB figures – this proportion was agreed in consultation with the Scottish Office.

Conclusion: The best estimate of bias is zero. The range for high and low variants is shown in Appendix H.

Appendix F: Asylum seekers /Visitor switchers

For both asylum seeker and visitor switcher data there are no sampling errors to estimate as the estimates come from 100% administrative data. The Home Office expert was consulted to help estimate sources of uncertainty and six other experts were consulted. The experts were: Keith Jackson (HO), John Hollis and Marian Storkey (LRC), John Salt (UCL), Garnett Compton and David Orr (GRO – Scotland) and Lucy Vickers (ONS).

Asylum seekers

People entering the UK to seek asylum are not likely to be interviewed by the International Passenger Survey and so asylum seekers are included in the mid-year population estimates using data provided by the Home Office. These data are in respect of the number of applications for asylum in the preceding mid-year to mid-year period. They include asylum seekers who are in detention. They are not net of refusals, although the Home Office does reduce the number of applicants by 10% to allow for potential asylum seekers who leave within a year of arrival. Other refusals would be covered by the IPS, as those people would be out-migrants on leaving the country. Therefore, any asylum seekers who remain here illegally are included in the population estimates.

Adjustments have been made for the years in the eighties before it was realised that a complete count of asylum seekers could not be achieved from the IPS.

The Home Office advice is that any error in asylum seeker data is negligible. There may be some uncertainty regarding what are referred to by the Home Office as special cases. These are where persons have been allowed to remain in the UK on an exceptional basis without formally applying for asylum. In the past, this was mainly people of Chinese ethnic origin who were refugees from Vietnam. The Home Office has made a provisional estimate of the number that might be involved of 11,600 but this may be an overestimate as some of these people were here before 1991 and it seems likely that they would have completed a Census form. The Home Office is currently reviewing this estimate of uncertainty.

Since early 1999 some refugees from former Yugoslav countries have been treated as special cases. The number of people in this category will be included with asylum seeker data for use in the compilation of future population estimates and thus will be included in 2001.

Conclusion: A possible bias of an underestimate of 11,600 on the population estimate. The Home Office is reviewing this estimate. If they confirm it as an official estimate of asylum seekers this number will be added in to a future population estimate. If they provide it as a possible source of uncertainty but not an official estimate then it shall remain as a source of uncertainty.

Visitor Switchers

A visitor switcher is a person who says they are entering the country as a visitor and then stays for a year or more. Estimates are made of non-European area nationals wishing to extend their stay in the UK.

There are possible sources of overestimation and underestimation in these figures. Numbers may be overstated because of the way in which visitor switchers are identified. A visitor switcher is identified as a person who has been granted an extension of stay that gives them a period of residence of a year or more but for whom there is no previous computer record. This is because visitors do not have computer records. A failure to link to an existing computer record would, therefore, generate a visitor switcher. There is no information about outward visitor switchers but it is thought likely that in practice many British nationals becoming visitor switchers abroad are likely to return home first to settle their affairs before the final emigration. They would therefore be covered by the IPS. Finally, the number of visitor switchers may be overstated if people who have an intention to stay for a year or more on entry tell immigration officers they will stay for less than 12 months but tell IPS interviewers

their real intentions. However, it seems more likely that if they were reluctant to tell immigration officers their real intentions they would also be reluctant to tell an IPS interviewer.

Possible sources of underestimation include the fact that there is no estimation of EEA visitor switchers. Free movement of EEA nationals means that statistics do not exist to allow assessment of visitor switching from EEA countries. It has been assumed that the net effect of visitor switching to and from EEA countries is zero. There is no estimation of illegal immigrants who enter the UK illegally. People who enter legally and stay on illegally are included in the estimates but people who avoid detection on entry cannot be included.

The current series of population estimates takes account of improved estimates of visitor switchers from 1981 onwards that the Home Office has produced.

Conclusion: On balance, Home Office advice is that figures for visitor switchers are probably correct, but a plausible range for these data could be up to +/- one-third. For 1981-98 data this represents an uncertainty range of +/- 105,400.

Appendix G Armed forces

The population of armed forces personnel and the dependants of foreign armed forces is estimated separately each year using the best data that are available at that time. If there were any errors in the data for a particular year these would not be compounded over time.

The Defence Analytical Services Agency provides annual data on the number of HM armed forces personnel in the UK by its constituent countries. These data are mainly based on administrative systems that are used for pay and personnel purposes. A snapshot is taken from the administrative records on the 1st of the month, this being a time that is linked to when pay runs are carried out. It is therefore DASA's view that the data are without any biases at the national level.

There may, however, be small errors in the armed forces component of the population estimates if the information on whether or not personnel are stationed overseas is not completely up to date in pay records. Each year DASA provides data for the number of personnel overseas and there is no reason to suggest that there would be large errors in these data. This point will be explored further with DASA.

As the armed forces component is estimated using fresh data each year, the errors are not cumulative. Therefore, in the mid-2001 population estimate, the only error resulting from the estimation of armed forces will be in respect of the mid-2001 armed forces data and that will need to be assessed with DASA at that time.

Conclusion: There is no sampling error as the data are based on 100% of administrative data. Assume there is no bias at the present time and reassess the position closer to 2001. The impact of the final uncertainty interval will be small as the size of the armed forces population is small. For instance, even if there were an allowance of +/- 10% for uncertainty this would only be +/- 20,000 on a population estimate in excess of 52 million.

Appendix H: Synthesis of uncertainty intervals

The table below shows how the estimates of uncertainty in the different components of the demographic estimates have been combined. It has been assumed that the different components are statistically independent. Thus the covariances are assumed to be zero and the variances can be summed. Biases have also been treated as additive – account was taken of whether the number is to added or subtracted from the census base. Thus for example the best estimate of the uncertainty due to possible bias in the IPS is that both immigration and emigration might be underestimated. This can be seen from the fact that the sign for the bias for each is negative - the net effect on the demographic estimate is that it should be increased by 25,834 (145,082-119,248).

In producing our best estimate of the range of uncertainty combining the results for all the individual components a view has to be taken on the net effect of the uncertainty of migration estimates. Immigration and emigration may each be either under or over estimated giving, in theory, four possible combinations for the net flows, i.e. both underestimated, both overestimated, immigration underestimated and emigration overestimated, immigration overestimated and immigration underestimated. If both flows are either under or overestimated the net effect is relatively small. If immigration is underestimated and emigration overestimated the net effect on the population would be to underestimate the population and the size of the underestimate could be quite large in absolute terms. Apart from the position for Eire, the most extreme combinations have not been presented in this synthesis because the views of the experts were that flows were either both underestimated or both overestimated.

The uncertainty interval combines the effects of standard error and bias on the estimate of total population, i.e. based on the table below:

$$\text{Minimum value} = 52,818,000 - 1.96(129,420) - (-33,711) = 52,598,048$$

$$\text{Maximum value} = 52,818,000 + 1.96(129,420) - (-33,711) = 53,105,374$$

$$\text{Half Range: } 254,000$$

One expert felt that it is possible that the IPS estimates for different years could be correlated rather than independent as assumed in our calculations. This is possible since the sample is clustered and more or less the same sample of PSUs is used each year, especially the case in the main air sample. However, because of the changes to the design, changes to airline schedules etc. these correlations are likely to be small.

It is not possible to calculate the extent of such potential correlation and so this has not been taken into account in the uncertainty intervals that have been calculated. The table below shows the possible impact of correlations:

Correlation	0.1	0.2	0.3
Increase in 20-year se	4%	8%	14%

If the IPS data for different years were correlated and assuming the extreme case where the standard error for the combined inflow and outflow were increased by 20%, this would result in an uncertainty interval of 52,547,316 to 53,156,106 (half range: 304,000).

Calculation of uncertainty estimates for 2001 rolled-forward population estimates (Persons)

England and Wales, projected 2001 population = 52,818,000
(1996-based population projections)

Component of Demographic rolled-forward estimate

	Bias estimates (best, low, high)				Count
	Standard deviation	Best estimate of bias ²	Low bias variant ²	High bias variant ²	
1981 census base (adjusted)	104,375	0	0	0	49,635,300
Births (20 years from 1981)	0	0	0 ⁵	0 ⁵	6,651,900
Deaths (20 years from 1981)	0	0	0 ⁵	0 ⁵	-5,758,200
IPS Immigration (20 years from 1981)	59,278	-145,082	-384,993	0	4,812,415
IPS Emigration (20 years from 1981)	48,388	-119,248	-244,596	0	-4,447,195
Immigration from Scotland and N Ireland & Isle of Man (20 years)	0	-61,093	-61,093	61,093	1,221,850
Emigration to Scotland and N Ireland & Isle of Man (20 years)	0	-64,816	-64,816	64,816	-1,296,322
Immigration from Eire (20 years)	0	0	-39,771	39,771	240,151
Emigration to Eire (20 years)	0	0	-40,457	40,457	-188,677
Visitor switchers/asylum seekers (20 years)	0	-11600	-117000	105400	514,000
Armed forces adjustment for 2001³	0				200,000 approx.
TOTAL	129,420	-33,711	-252,988	100,991	(52,818,000)⁴

Notes:

1. variance includes sampling and non-sampling variance
2. Biases are expressed as numbers to be subtracted from the population estimates, (i.e. a negative bias Means that the estimate is too low)
3. Armed forces figures for 2001 will be used - the potential uncertainty in the figures will need to be Assessed nearer the time
4. Projected population for England and Wales in 2001 is 52,818,000. In 2001 this column will be based on the rolled forward population estimate that is produced using the component method and so will sum to the point estimate of the population exactly. As a projection for 2001 is being used at the present time this column does not sum to the point estimate.
5. It has been assumed at the current time that there is no bias in the birth and death registration based data. This will be reviewed closer to 2001 although given the high quality of these data it is unlikely that an uncertainty interval would be wide.

Appendix I: The Delphi study of international migration

Delphi studies are a means of eliciting expert opinion and achieving consensus. The technique provides a method for structuring communication in a way that allows a group of respondents to confront a complex problem and to reach consensus¹. In each round of the study a questionnaire is distributed to the members of the respondent group. Completed questionnaires are returned to the co-ordinating group who prepare a summary of the results for inclusion as a preamble to the following rounds. This continues until the stability of responses appears and general convergence of opinion emerges, or a respondent group fatigue factor emerges, or a point of diminishing returns is reached. The questionnaires are filled in confidentially by each member of the respondent group without consulting other participants. The technique has been used successfully in the UK to reach agreement on sensitive budgetary issues².

To obtain opinions about possible non-sampling variance and biases in the IPS the ONS sent out the questionnaire below to the following experts who participated in the study: John Salt, David Orr, Garnet Compton, Ludi Simpson, and from ONS Lucy Vickers, Dave Elliot, Peter Lynn, Tim Jones and Iain MacLeay.

Once the questionnaires had been returned the results were summarised. The summary and an anonymised list of comments were sent back to respondents who were then invited to reconsider their responses. There was some convergence of opinion by the end of this second round, but some respondents did not change their minds. The responses to the second questionnaire are those used for estimating uncertainty. A weighted average was calculated, based on two scores: 1) for respondent's experience and 2) for the strength of evidence presented by the respondent. Because there was some variance in the results a high and low variant were also calculated (see Appendix C).

The Questionnaire

This questionnaire covers only IPS flows and migration flows between the constituent countries of the UK. If you want to provide expert comments on the other components please contact John Charlton at john.charlton@ons.gov.uk (0171 533 6239) for a questionnaire.

International Migration data

Migration figures are given for you to use as a guide when completing this questionnaire. Table 1 shows the IPS figures for the ten-year period 1981-1991

Column (a) gives the estimate of international in-migration

Column (b) gives twice the standard error, from which 95% confidence intervals can be constructed.

Columns (c) and (d) give the same information for international out-migration

Columns (e) and (f) give the balance, i.e. net international migration and twice its standard error.

We have provided a list of potential sources of non-sampling error. These are given on page 4. Furthermore, as a starting point, a suggested range for errors and biases is given on the questionnaire based on the opinions of a few experts, as an aid to starting the process, but you should not feel unduly influenced by these suggestions.

The purpose of this exercise is to produce a plausible range for the Census 2001 estimate of population at the national level. These opinions cannot be regarded as official estimates of possible bias and non-sampling error on international migration data. The survey figures should be regarded as confidential and should not be quoted or used in publication.

The suggested errors are for total persons (all ages, both sexes). We are asking you to give your estimates of the possible size of the errors and biases. A point estimate, either a percentage or a number, would be preferred, but a range would be acceptable. Ultimately ONS will need to apportion these by age and sex. In the absence of any

better information this will be done equally across the age and sex groups. If you have any knowledge about specific age or sex groups that would help us achieve this in a better way please provide this information.

Non-sampling errors can affect results in two ways - they can increase total variance, and they can cause a bias. Because these affect our calculations in different ways we are asking you to quantify these separately. For example, with bias such as a 5% undercount each year, there will be a 5% undercount over 20 years. If non-sampling error results in an effective increase in the standard error from **S** to **S*X**, then over twice the time period the standard error is would be around $1.4*S*X$.

Your estimates of the bias and effect of non-sampling error

Notes

- 1. Please do not try to account for asylum seekers and visitor switchers because these are being dealt with in a separate questionnaire. Migration between the UK and Ireland is also being dealt with elsewhere
- 2. We are interpreting both the “%bias” and the “%effect on standard error” as the (estimate – true value)/estimate. Thus a negative value indicates that the survey produces an under-estimate and a positive value indicates an over-estimate.

In-migrants

Give one line of figures for total flow for all ages, both sexes, E&W

Bias	Effect on standard error
<i>(% of flow)</i>	<i>(% of standard error)</i>

Your suggested estimate

Out-migrants

Give one line of figures for total in flow all ages, both sexes, E&W

Bias	Effect on standard error
<i>(% of flow)</i>	<i>(% of standard error)</i>

Your suggested estimate

Apportionment by age and sex

Do you have any specific knowledge to help us apportion non-sampling errors by age and sex? If so please provide the information below or on a separate sheet.

Finally

Please indicate on a scale of 0 to 10 the degree of confidence you have in your opinion, with 0 indicating “no confidence at all” and 10 “I speak with total authority”. This may be used in arriving at an overall figure. []

If you have any further comments you on international migration that might be helpful please write these below or on a separate page.

Table 1

**INTERNATIONAL MIGRATION
MID YEAR 1981 TO MID YEAR 1991**

ALL PERSONS

	INFLOW		OUTFLOW		BALANCE	
	GROSSED	2x Standard Error	GROSSED	2x Standard Error	GROSSED	2x Standard Error
ALL AGES	2,383,405	65,766	2,332,703	57,938	50,702	87,646
UNDER 15	464,777	28,322	439,165	24,244	25,613	37,280
15 - 24	720,968	37,890	571,048	29,832	149,920	48,224
25 - 44	977,429	42,016	1,053,199	38,582	-75,771	57,044
45 - 64	184,173	16,420	211,085	18,136	-26,912	24,466
65 & OVER	36,058	7,240	58,206	7,840	-22,148	10,672
MALES						
ALL AGES	1,168,735	46,238	1,216,362	41,680	-47,627	62,252
UNDER 15	239,362	20,688	219,013	17,138	20,350	26,864
15 - 24	306,385	25,334	262,396	20,222	43,989	32,416
25 - 44	498,612	29,736	590,016	28,758	-91,404	41,368
45 - 64	104,622	12,238	119,000	13,482	-14,378	18,208
65 & OVER	19,754	5,842	25,937	5,082	-6,183	7,742
FEMALES						
ALL AGES	1,214,670	46,768	1,116,341	40,242	98,329	61,698
UNDER 15	225,415	19,342	220,152	17,148	5,263	25,848
15 - 24	414,583	28,176	308,652	21,932	105,931	35,706
25 - 44	478,816	29,684	463,183	25,720	15,633	39,278
45 - 64	79,551	10,948	92,085	12,132	-12,534	16,342
65 & OVER	16,304	4,278	32,268	5,970	-15,964	7,344

NOTES

Potential Sources of Non-Sampling Error in International Passenger Survey (IPS) Migration Figures

This note aims to suggest areas where IPS non-sampling errors may impact on the accuracy of migration estimates.

Coverage

- No IPS interviews are normally conducted at the smaller airports (e.g. Belfast International, Southampton, Liverpool, etc) or on sea routes with small volumes of traffic.
- IPS interviews are only conducted between 6:00am or 7:00am and 22:30pm. Allowances are made for “out of hours” traffic, but migrants weights based upon interviews conducted. Some long haul flights to the UK land before the IPS start conducting interviews.

Non-response error

- The IPS response rate is normally about 85% for complete and partial interviews.
- Non-response may occur because the contact does not speak English; is in a rush to go duty free shopping (departures) or to meet family or friends etc (arrivals); engaged in illegal activity; or just does not wish to co-operate with Government surveys.

Interviewer error

- The interviewer misunderstands the response from the contact.

Respondent error

- The respondent accidentally misinterprets the question.
- The respondent, mistakenly believing that there is a link between IPS and immigration or customs, deliberately provides false information.
- The respondent believes they will stay more than a year at the start of their visit (classified as a migrant), but change their mind and leave within the year (classified as a visitor).

Grossing numbers are estimated

- IPS grossing totals are estimated in some cases. The number of passengers on Channel Tunnel shuttles are estimated based upon the number of vehicles transported, provided by Eurotunnel, and the average occupancy rates of vehicles, estimated by the ONS’s vehicle occupancy survey.
- Air passengers are grossed up based on traffic volumes supplied by the CAA less those passengers who “interline”. Interliners are defined as passengers who just change planes in the UK before continuing their journey and who pass through the UK without passing through passport control. The number of interliners is estimated from a CAA survey.

Data handling errors

- The questionnaire response is either incorrectly written on the form, or the results are incorrectly typed into the CADI system (computer aided data input).
- The IPS processing system miscodes information.
- Any grossing methodology may distort migration numbers. ONS uses a complicated grossing method to minimise the chance of distortion. For example, for air traffic at the main airports (Heathrow, Gatwick and Manchester), ONS takes the initial sampling probability, uplifts this due to non-contacts/refusals, uplifts for minimum responses, uplifts for out of hours traffic, and adjusts the results to the traffic volumes. A similar procedure is used for contacts at other airports, but these contacts are grossed up to volumes including those at the smaller un-sampled airports. Similar methods are also used for sea and channel tunnel traffic.

IPS: Some further Notes given out at Round 2

Out of Hours Traffic – Current Practice

Within the IPS assumption are made that the proportion of in-migrants and out-migrants is the same in the omitted parts of the population as in the rest of the population. However, this assumption is not made on a global basis, but is made for smaller sub-sections.

Contacts at each terminal at Heathrow, Gatwick and Manchester are grouped into 7 regions of the world by flight origin or destination. The out of hours traffic is grouped into the same 7 regions. The contacts are then weighted within each region to account for the out of hours traffic. Irish Residents are excluded from this stage of the weighting. The 7 regions are shown in figure 1 below.

Figure 1. Regions used in out of hours weighting

Group	
1	Europe (including all former USSR)
2	Africa
3	Pacific (incl. Australia and New Zealand) and Asia (Exc. Hong Kong)
4	Middle East, South America and other
5	Japan
6	Hong Kong
7	North America

The following tables details the out of hours passenger volumes at the 8 major air terminals for the 7 regions, for one specific quarter, for arrivals:

Table 1. (thousand arrivals)

	A	B	C	D	E	F	G	Total
Heathrow 1	20	0	0	11	0	0	1	32
Heathrow 2	18	2	0	13	0	11	1	45
Heathrow 3	10	24	8	1	0	39	4	85
Heathrow 4	20	35	85	1	0	48	2	190
Gatwick South	264	23	9	22	0	0	23	341
Gatwick North	104	4	2	48	0	0	24	182
Manchester T1	208	22	1	7	0	0	6	244
Manchester T2	159	7	2	4	0	0	10	183
Total	803	118	108	107	0	97	70	1303

A similar table can be constructed for departures etc.

At the smaller residual airports, the implicit assumption is made of there being similarity between the omitted part of the travelling population and the rest of the population.

Current IPS Development Projects

Out of hours traffic

A research exercise will probably commence, at some of the major terminals, in 1999 to test the hypothesis that the characteristics of passengers travelling out of hours are similar to those travelling in the standard periods. The results of this exercise, however, will not be available in time to feed into this study of migration statistics.

Exclusion of small airports

An assumption is also made within the IPS that the characteristics of passengers at small airports are similar to those at the “residual” airports. All sampled airports, with the exception of Heathrow, Gatwick and Manchester airports are currently used as the proxy for these smaller airports.

A study was started in 1998 to again test this hypothesis. IPS shifts were run throughout 1998 at a number of the smaller international airports (Liverpool, Southampton and Bournemouth). These results will be analysed in 1999 to see if the characteristics of travellers (reason for visit, spending patterns etc) are similar to those found at the residual airports (e.g. Bristol, Glasgow, Birmingham etc) which are currently used as the proxies for this traffic.

Exclusion of small volume sea routes

A similar test is being carried out to examine some of the smaller volume sea routes. In 1998 interviews are being conducted on the small volume routes, e.g. between Newcastle and Amsterdam, Gothenburg and Hamburg, Portsmouth and Santander etc. Currently traffic on these routes is grouped with that from similar routes.

Non-response bias

Attempts are made to gain as much information from every contact possible. Sometimes only a minimum level of information is obtained, such as nationality or residence. This information is used in the weighting process, with the weights of these contacts “transferred” to similar records based initially on flow, purpose and nationality or residence.

It is unfortunately not usually possible to use luggage tag information, since the interviews are conducted before arrivals pick up their bags and after departing passengers have checked in their luggage. The new “maroon” EC passports also make it less easy to determine the nationality of a contact from just seeing the outside of their passport.

From the start of 1999 the IPS is trying to get a better handle on language problems by asking the interviewer to note down the language of the contact – if it is believed that a refusal, non-contact, partial or minimum response was due to language difficulties.