An international comparative study on the use of the Cohort Component Method for estimating national populations

Nicholas Ormiston-Smith, Jonathan Smith and Alison Whitworth
Office for National Statistics

Introduction

The Office for National Statistics (ONS) produces the Registrar General's population estimates and projections for national and subnational populations in England and Wales. This report compares the process used to apply the Cohort Component Method (CCM) in England and Wales with the processes used in other countries that use CCM outside the UK. The aim of this article is to identify how other countries calculate their population estimates and briefly consider their application in England and Wales. ONS may be able to adapt these alternative approaches to improve national estimates. No testing of alternative approaches has been conducted as part of this review but, where it would be effective to do so, there will be further research into potential improvements. This work would be undertaken as part of ONS's Improving Migration and Population Statistics (IMPS) project. Further details about this project can be found at www.statistics.gov.uk/imps.

Much of the information in this study and in particular information on how National Statistical Institutes (NSIs) apply the CCM to estimate their national populations has been collected through the use of a questionnaire and follow-up correspondence. Countries with good quality birth and death registration data are the primary focus of this investigation as they are most relevant to England and Wales. Where published sources have been used this will be referenced in the text.

CCM is a method of rolling forward age and sex specific population estimates from one year to the next. In brief, births are added to the population, deaths (by age group or single year of age) are subtracted, the population is aged a year, and an adjustment is made for net migration (again by age group or single year of age). CCM is widely used by NSIs around the world for both population estimation and projection.

This comparative study explores the use of the Cohort Component Method (CCM) to produce national level population estimates. This method is used annually to calculate mid-year population estimates for England and Wales by the Office for National Statistics (ONS).

Initially the article considers recent population change in England and Wales, with particular emphasis on the growing importance and challenges faced by migration estimation. Comparisons are then made between how population estimates are produced in England and Wales and other countries, with a particular focus on differences in the way the CCM is applied. Recent changes in methods used to estimate population are then reviewed along with a discussion of alternative approaches such as those described in academic literature.

The key difference between projections and population estimates is that projections are based on assumptions about future trends, based on age-specific rates. Population estimates are based on information from past events, that is, mid-year estimates for 2005 were published in August 2006. If data used to estimate the components of population change are of high quality, the cohort component method will be the best method for measuring change in the total population.

In England and Wales, after the 2001 Census there was a gap between the mid 2001 estimate rolled forward from the 1991 Census and the equivalent estimate based on the 2001 Census. This was partly due to the rolled forward estimates overestimating the population. As birth and death registration is known to a high degree of accuracy then migration estimates are a large cause of the discrepancy. However, another reason for the discrepancy is definitional issues. Migration is a major component of population change in England and Wales. In addition, the nature of migration has changed over recent years which has increased the difficulties in measuring migration.

In this article a distinction is made between national and subnational population estimates. In subnational population estimates, factors such as subnational and international migration distribution need to be taken into account. These factors do not affect population change at the national level.

The article will first describe how population estimates are used. Consideration is then given to the components of population change measured in the CCM including background as to how these components have changed in recent years. Given this context, we review how the CCM is applied to produce national population estimates for England and Wales and examine differences in the application of the method in other countries. Finally, recent improvements in other countries methods and other methods of estimating populations are considered.

Key areas of discussion include:

- resident definition – either *usually resident* population – all the people who usually live in a given area at a given time, or *de facto* population – all the people who are currently present in a given area at a given time
- population base – from which population estimates are rolled forward, usually from a census or population register
- data sources – information sources used to compile population estimates
- migration – data sources and methods of measurement
- Adjustments – in particular those used for migrant or visitor switchers (defined later)
- age groups – different data sources may be available for different age groups
- static populations – populations which have fairly constant age distributions over time such as armed forces or students, particularly where these groups are not measured through conventional measures
- quality – methods used in the assessment of quality

Use of population estimates

The primary uses of population estimates in England and Wales are set out in *Making a Population Estimate in England and Wales.*[1]

Mid-year population estimates currently have a wide variety of uses within central government, as well as being used by local authorities and health bodies, other public bodies, commercial companies and individuals in the private and academic sector. These uses can be categorised into two broad groups:

- Uses where the absolute numbers are of key importance. This may be in terms of allocating financial resources from central government, planning services or grossing up survey results. Some of the main central government uses are concerned with resource allocation and are carried out by the Office of the Deputy Prime Minister [now Department for Communities and Local Government] for England, and by the Welsh Assembly Government.
- Uses where the population figures are compared with other figures such as the numbers of births or deaths in the calculation of rates and ratios.

Population estimates are also used by the Boundary Commission for England and the Boundary Commission for Wales when assessing electoral boundary revisions. Estimates have political importance around the world. For example in Australia 'The Commonwealth Electoral Act 1918 requires the Statistician to supply all such population statistics as requested by the Australian Electoral Commission for the regular review of the number of seats each State is entitled to have in the House of Representatives'.[2] In a European Union (EU) context national population estimates are used for 'Qualified Majority Voting' a voting procedure employed in the Council of the EU for taking some decisions. Regional population estimates are used for the allocation of EU funds at a regional level.

Population estimates, wherever they are produced and whichever method is used to produce them, are required to be as accurate as possible given the importance of what they are used for. The implication of this for the CCM is that the components of population change are required to be measured as accurately as possible.

Components of population change

This section describes the components used to measure population and how the pattern of one of those components, migration – the most difficult component to measure, has changed in recent years.

The accuracy of the CCM is dependent on the quality of data available to measure components of population change (births, deaths and migration). Births and deaths are measured using registration data. As registration is a legal requirement, the data for the measurement of natural increase (births *minus* deaths) is from a high quality source and is therefore thought to be very accurate. There are some known difficulties such as the use of appropriate residency definitions.

Measuring migration has always been more difficult partly because, unlike for births and deaths, there is no compulsory system of migration registration in England and Wales. Also large changes in the number of migrants are more common than for births or deaths. Recently, the measurement of migration has become more critical because the nature of population growth in England and Wales has changed. Natural increase used to be the major driver of population growth, even though there was net emigration in 1992 and 1993 the population increased by 127,000. Figure 1 shows how the migration proportion of population growth has increased since 1991. Changing patterns in international migration have been explored further in *Focus on People and Migration*.[3]

Figure 2 shows that while both inflows and outflows are increasing, inflows are increasing at a greater rate than outflows therefore net migration is increasing.

Figure 3 shows that the biggest difference between inflows and outflows are for males and females aged between 15 and 24. There are also slightly more people aged between 25 and 44 migrating to the UK than emigrating from the UK.

It has become increasingly difficult to measure migration due to the increasing amount of foreign travel and the increasing number of significant arrival and departure points.[4] Table 1 shows how international traffic has increased at regional airports from 1980 to 2004. Between 1980 and 2004 the increase in passenger numbers at London airports was approximately three-fold. During the same time period the increase in passengers at regional airports was almost seven-fold.

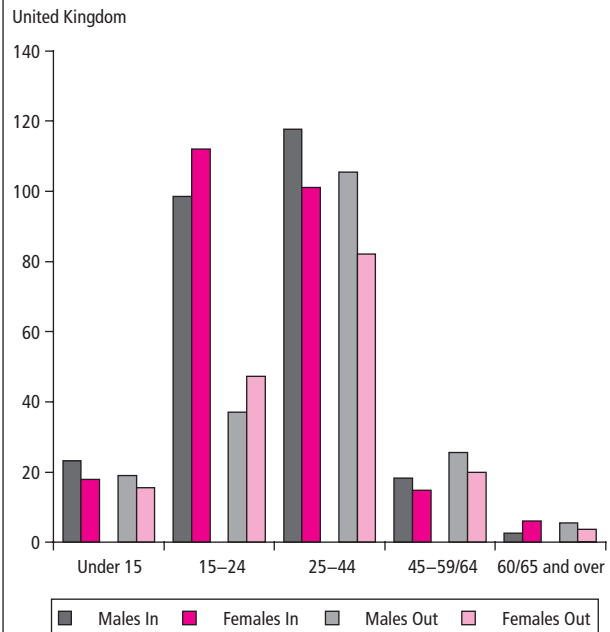| Figure 1 | Population growth in England and Wales by migration and natural increase from 1991–2004 |
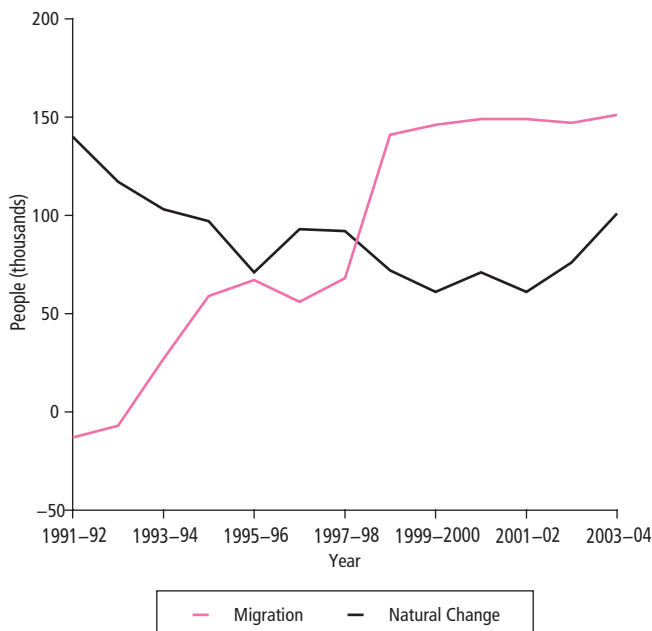
Source: Population Estimates, ONS

| Figure 2 | International migration flows, 1991–2004 |

United Kingdom

Source: International Migration, ONS

| Figure 3 | International migration: by age and sex, 2003 |

United Kingdom

Source: International Migration, ONS

| Table 1 | Total international traffic at UK airports, 1980–2004 |

	Passengers (millions)			
	1980	1990	2000	2004
London airports	34.5	58.7	102.1	112.3
Regional airports	8.4	18.6	40.6	54.0
Total	42.8	77.4	142.7	166.3
Regional share of total (per cent)	*20*	*24*	*28*	*32*

Note: 2004 statistics are for the 12 months ending November.

Source: CAA Airport Statistics, terminal passengers at all reporting UK airports (excludes Channel Islands)

A further difficulty in counting the number of migrants is the small number of migrants in comparison with the total number of travellers. Approximately 1 per cent of all travellers since 1991 are migrants – shown in Figure 4. The International Passenger Survey (IPS) is the main source used to measure international migration in England and Wales. The IPS is a sample survey carried out by ONS for a range of public and private sector organisations. In particular, the survey provides figures used for the travel account of the balance of payments and for informing decisions on tourism policy in addition to international migration statistics. The ONS Centre for Demography funds migration filter shifts in order to increase the sample size of immigrants. During these shifts a full interview is only given to those people who are intending on staying in the UK for twelve months or more.

Figure 4 | **Number of people arriving in the UK: 1991–2004**

United Kingdom

Source: International Passenger Survey, ONS
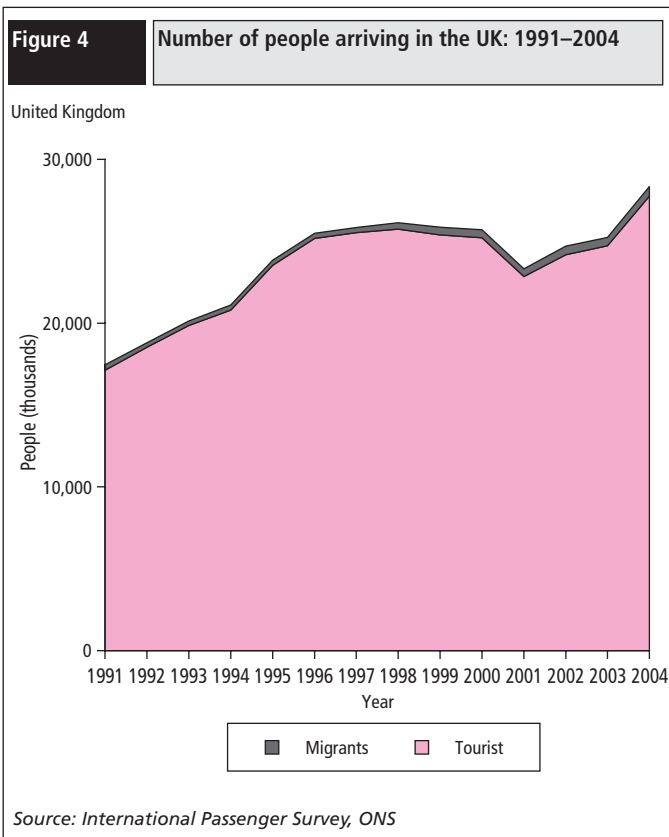
Calculating population estimates

Population estimates for England and Wales are calculated using CCM. This can be summarised by:[1]

1. Take the previous mid-year resident population (or census if taken in the previous year) and age-on by one year (or for the period between census and mid-year).
2. Then estimate 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 population of the area.

In England and Wales the population estimates are based on the usually resident population, that is, based on where people usually live rather than where they happen to be staying on Census night (see Box One for a full definition). The population base is a decennial census, estimates are rolled forward annually from this base. Population estimates are rebased using the new census results when those become available. An adjustment is made within the census process to account for undercount in the census, known as One Number Census (ONC) for the 2001 Census.

The definition for resident population in England and Wales differs from that used by Eurostat which states that the resident population includes armed forces and diplomats stationed abroad but excludes foreign armed forces and foreign diplomats. Although many countries use the Eurostat definition, in England and Wales the definition includes foreign armed forces and excludes UK armed forces and diplomats overseas.

There is a demand from users, such as local authorities, for population estimates based on different population bases. For example, workday population or weekend population. ONS is currently researching methods for providing estimates using alternative definitions. For more information see *Population bases and statistical provision: towards a more flexible future?*.[5] Alternative definitions present new challenges for the

Box one

Definition of Usual Resident

The population base from the 2001 Census underpins the mid-year population estimates resident base and is defined as follows:

> The 2001 Census has been conducted on a resident basis. This means the statistics relate to where people usually live, as opposed to where they are on Census night. Students and schoolchildren studying away from home are counted as resident at their term-time address. As in 1981 and 1991, residents absent from home on Census night were required to be included on the Census form at their usual/resident address. Wholly absent households were legally required to complete a Census form on their return. No information is provided on people present but not usually resident.

From *Making a Population Estimate in England and Wales* (Jefferies and Fulton 2005)

CCM method. For example what sources would be used to roll forward estimates of the workday population?

In other NSI's population estimates are calculated using two main methods. The first uses CCM as in England and Wales, the second is only available to countries that have a Population Register who are able to produce population estimates from it. A universal population register is a government data collection system in which the demographic and socioeconomic characteristics of the entire population are continuously recorded. Registers can be used for demographic purposes because they record the major demographic events (birth, marriage, migration, death) and so up-to-date information on the whole population is readily available.

The EU-15 countries with a population register that base their population data entirely on a register are Austria, Belgium, Denmark, Finland, Netherlands and Sweden. Two countries, Italy and Germany, have a population register, but use a component method to estimate the population. A census provides the population base and intercensal data are estimated on the basis of the population changes in the population register. Tables 2 and 3 summarise the sources available and the methods used in the EU-15.

A key advantage of a population register is the population coverage it can attain. Individuals are added to the register at birth and in principle the population for any particular day can be determined. One area where registers may be prone to error is migration, even though people are required to notify authorities on arrival/departure. It is unlikely that the population register will have complete coverage of international migration movements and so countries with population registers may have similar issues as countries without registers in this respect.

Table 2 | **Sources available for population estimates**

Centralised population register	Decentralised population register	No population register
Belgium	Germany	Greece
Denmark	Spain	France
Finland	Italy	Ireland
Sweden	Netherlands	Portugal
Luxembourg	Austria	United Kingdom

Source: Final summary report concerning the methodology of regional population estimates

Table 3	Methods used to create population estimates

Population register	Census and population register	Census and civil register	Census and projection
Belgium	Germany	Greece	Spain
Denmark	Italy	Ireland	
Finland		Portugal	
Netherlands		France	
Sweden		UK	
Austria			

Source: Final summary report concerning the methodology of regional population estimates

Scandinavian countries notify each other if a resident of one country becomes resident in another. This information enters into the population estimates with a time lag because different time criteria are used in each country. For more information on time criteria and other differences see Poulain *et al* 2006.[7] The error related to emigration may increase as movement around the EU becomes easier.

Most countries produce estimates on a usually resident definition, as is the case for England and Wales. Ireland is one of the few countries that for historical reasons uses a de facto population. They do however intend to move to a usually resident definition in the future when an EU regulation on migration statistics comes into force, as this requires a usual residence definition.

The reference date of population estimates in many countries is either 31 December or 1 January. Population estimates made for the UK and Ireland refer to 30 June and 15 April respectively. In England and Wales these are known as mid-year population estimates. For many demographic statistics such as fertility rates, a mid-year estimate is preferred. Although there is a seasonal dimension to population estimates in the UK this is more likely to be important at a subnational level. Defining usual residence becomes more complex when migrants spend equal amounts of time either in different areas of the same country or between different countries. ONS is undertaking separate research into population definitions.

Most countries that use CCM for their population estimates use a census as the population base because it is assumed to be the most accurate estimate of the population. In Lithuania population census data are used and updated by natural increase/decrease and net migration is based on a residents' register. In Latvia annual population estimates are based on a population register (since 2000) and legal residence is considered usual residence.

Data sources

The accuracy of the CCM is dependent on the quality of data available to measure components of population change. In England and Wales, as already noted, births and deaths are measured using high quality registration data. As registration is a legal requirement, the data for the measurement of natural increase (births *minus* deaths) is from a high quality source and is therefore thought to be very accurate though not perfect. Estimates of the population under one are being considered as part of the IMPS project. Migration is measured using a number of sources but the principal data source for international migration is the IPS. Approximately 250,000 interviews are carried out per year representing 0.2 per cent of all travellers as they enter or leave the UK. Most of these are visitors, approximately one per cent of the sample are likely to be migrants.

Home Office data on asylum seekers and their dependants are also used in calculating international migration. Information from the Irish Quarterly National Household Survey is used for flows between England & Wales and the Republic of Ireland.

Other sources of data for mid year population estimates include:[1]

- The Defence Analytical Services Agency – used for armed forces
- Information from United States armed forces –- used for foreign armed forces
- Department for Education and Skills and Welsh Assembly Government – used for school boarders
- Home Office data – used for prisoners
- GP registrations – used for internal migration.

In the previous section we have seen there are a number of countries which can estimate their population using a population register. If however a country doesn't have a register they must create the population estimates from a number of sources. In countries without a population register a census will be the most accurate population base (usually taken every five or ten years). The measurement of births and deaths will normally be based on a vital statistics registration system. The registration system will have almost complete coverage and the quality will be extremely high. The measurement of migration is more difficult and each country will have a different data source which can provide the highest quality estimation of migration. Some examples of these alternative data sources are an administrative data source such as passenger card data in New Zealand, a survey like the Labour Force Survey which provide international migration estimates in Ireland or a registration system such as that used in Italy.

Administrative sources can often be used as a proxy to measure migration. The quality of the estimates will depend on the primary purpose of collecting the data. In the Czech Republic data on the immigration and emigration of foreigners is obtained from Immigration Police registers, so the coverage of legal migration should be comprehensive but measurement of illegal migration will be poor, as in most other countries. In addition, it does not cover migration of nationals. Table 4 gives examples of countries with different types of data sources for measuring migration.

In *Demographic Statistics: Definitions and methods of collection in 31 European countries*,[8] 19 of the 31 countries under study have a registration system for migrants. Registration data are a subset of administrative sources. A source is considered to be registration data if the primary purpose of collecting the data is registration rather than another administrative reason. For example, passenger cards are collected for immigration purposes and the extraction of migration information is a secondary use of the data. Where registration systems for migration exist they will have good coverage but are unlikely to be as complete as birth and death registrations. They may be more accurate if it is a legal requirement and if it is easy to comply, that is, data collected at point of entry. Immigration is likely to have better coverage than emigration as there is more incentive to register in a country when you are arriving, for example to enable access to health services, than to de-register from a country when you leave.

The length of time that a new entrant can live in the country without a requirement to register differs for different countries. In Sweden and Finland the time requirement is one year which relates to the UN definition of migration. In Denmark and Japan, however, the time limit is three months. This can cause difficulty in differentiating between short-term and long-term migration within countries and when making international comparisons. The proposed EU legislation on migration

Table 4	Type of data source for migration

Migration Source	Country
Administrative	France
	New Zealand
	Australia
	Canada
	Czech Republic
	Latvia
Survey	Ireland
	United States
Registration	Estonia
	Czech
	Romania
	Hungary
	Italy
Administration and Survey	England and Wales
	Portugal
	Cyprus
Census	Bulgaria

Source: Compiled from various sources

statistics, currently under review by the European Parliament and the Council of the EU, attempts to improve the harmonisation of migration statistics by introducing a common twelve month definition of migration.

Surveys, used to estimate international migrants, maybe prone to error from a number of sources. As with any survey there maybe sampling and measurement error. This may be exacerbated, as with the IPS, because migrants are a small proportion of the population being sampled. Bias can be introduced into estimates based on international migration surveys if respondents have a motive for not responding to the survey. Illegal migrants maybe suspicious of authority so less willing to respond to the survey. Finally, surveys maybe based on intentions, as in the passenger survey in Cyprus. This can lead to visitor switching which will be discussed later in the article.

On occasion a mixture of sources can provide superior estimates if a high quality source only covers immigrants and not emigrants or vice versa. In Cyprus immigration and emigration are both collected from a passenger survey, carried out at the point of arrival and departure. Information from expiry of residence permits for foreigners is also available for calculating emigration.

In Bulgaria their census is used to create migration estimates retrospectively because of a lack of alternative sources. There are several reasons this would not be appropriate in England and Wales. Migration figures are required annually, it would not be possible to wait until the next census. In addition migration statistics from a census do not indicate how patterns have changed since the previous census so it would be unclear what migration flows were annually.

Migration

Migration, as the component of change that is most difficult to estimate, has been discussed in previous sections but the issues are summarised here. Migration is more difficult to measure than natural change particularly where there is no system of registration. Where registration systems for migration do exist they will have good coverage but are still unlikely to be as complete as birth and death registrations, particularly with respect to emigration. International travel and migration is now much easier than in the past, particularly within the EU. There has been an increase in the number people who are travelling in and out from a greater range of ports/airports.

However data are collected, the definition of how an international migrant is defined is always relevant to the estimation of population. In England and Wales an international migrant is defined as someone

Table 5	Duration of residence required by migrant

Duration of residence required	Country
12 months	Australia
	England and Wales
	Estonia
	Finland
	France
	Greece
	Iceland
	Ireland
	Latvia
	New Zealand
	Portugal
	Romania
	Sweden
6 months	Norway
	Lithuania
3 months	Czech
	Denmark*
	Japan
2 months	USA
Intention of residence only	Austria
	Italy
	Luxembourg
	Spain
	Switzerland

* Residence required in Denmark is six months for other Nordic nationals

Source: Compiled from various sources

who changes his or her country of usual residence for a period of at least a year, so that the country of destination becomes the country of usual residence. This is also the case for most other countries but there are some exceptions (see Table 5). These examples occur when the data source used to measure migration does not use a one-year migrant definition. For example, in the Czech Republic migration data are based on applications for visas which are needed for stays over 90 days. The USA uses the American Community Survey (ACS) to measure migration and the ACS has a two month residence rule for inclusion in the survey, so any international migrant who has been in the USA for at least two months is eligible for inclusion in the data on which the international migration estimates are based. It can be assumed that when data sources are used that are unable to count people resident in the country for at least twelve months it is because there is no data source available that can identify those people. This may lead to an over-estimation of migrants (according to the UN definition[9]), a three month resident definition will suggest more migrants than a twelve month definition. In England and Wales IPS interviews are carried out with people whether they plan to stay for three to six months, six to eleven months or greater than twelve months. ONS is undertaking work researching short-term migrants.

Australia and New Zealand have some advantages when it comes to measuring migration. Their geographical isolation makes it very hard for anyone to enter the country without passing through official channels. This means there is virtually 100 per cent coverage of international migration in these countries. England and Wales share some of the same advantages by being an island but there is still a land border with Scotland and the short distance to continental Europe means flows are likely to be considerably higher. In contrast, Germany shares land borders with nine other countries.

Changes to immigration rules regarding EU partners mean that a large section of the potential immigrant population do not need to have a visa before working in England and Wales. This is a key problem with the

use of visas for migration measurement. Some countries in continental Europe face further challenges in the estimation of international migration. The Schengen Convention[10] abolished checks at internal borders of the signatory countries meaning there is free-flowing traffic across borders.

Visitor switchers

In England and Wales an adjustment is made for 'migrants switchers' and 'visitor switchers'; people who change their intentions and therefore their migration status. A visitor switcher therefore is someone who intended to stay in the country for less than 12 months but has subsequently decided to stay permanently. Estimates of immigration would be too low without taking visitor switchers into account. A migrant switcher is someone who intended to migrate permanently but has subsequently left the country within 12 months. In this case the estimate of migration will be too high. Since the mid-2003 population estimates, the migrant switchers adjustment has been made to the IPS data at source and therefore switchers are not estimated separately within population estimates. Since 2001, visitor switchers have been estimated from IPS data on two categories of visitor: those who initially intend to stay for 6–11 months, and those who state that they may stay for longer than a year although intended length of stay is uncertain. Prior to 2001, the Home Office provided data on visitor switcher inflows.

Australia and England and Wales are the only countries from which information was collected, to make an adjustment for visitor switchers – known as category jumpers in Australia. New Zealand is aware of visitor switchers but don't make an adjustment as they believe its effect to be small. Instead revisions are made to the population estimates after a census. The adjustments made in New Zealand may be smaller than England and Wales because their census is undertaken every five years, rather than every ten. The Statistics New Zealand webpage which covers population estimates states; "Our best information on the current net effect of category jumping is that it is negligible. Over the past quarter-century category jumping has been estimated to have varied between 0 and approximately 5,000 per annum; thus the intercensal effect has been in the range 0–25,000, or 0 – 0.8 per cent of the population."[11]

Australia, unlike the UK, is able to identify category jumpers from the data available and therefore they are able to make precise adjustments to the population estimates at small cost. The data available to the Australian Bureau of Statistics (ABS) for this purpose are passenger card data (where travellers express their trip length intentions) and passport and visa data available from Department of Immigration Multicultural Affairs.

The Australian method involves delaying the compilation of quarterly population estimates until an extra quarter of overseas migration data becomes available [and matching the passenger card data from arrivals and departures to check actual movement corresponds with intention]. This delay enables the accuracy of preliminary category jumping estimates to be improved significantly. "Estimates of the percentage of (intended) short-term Australian residents who will not return within six months, and the percentage of (intended) short-term overseas visitors who will stay more than twelve months are made, based on the experience of recent years for the quarters not yet available. The advantage of this method over previous ones used is that it can take account of sudden changes in category jumping patterns."[2]

In England and Wales comprehensive passenger card data on inflows and outflows are not available. A method previously used by ABS to estimate rates of visitor/migration switchers, by country of birth, applied to estimates of visitors arriving and residents departing could be used by ONS if suitable data on visitor/migrant switchers can be collected. This would require comprehensive data on inflows and outflows, linked with

data on intended/permission for length of stay on arrival. Comprehensive inflows and outflows may be obtainable by 2014 via the e-Borders Programme (see Box Two towards the end of the article) but further research on how to account for switchers would require further research. Visitor/migration switcher rates may be relatively stable over time but are subject to error if circumstances change, such as immigration policy of a particular country changes. New questions were introduced to the IPS in 2004 and 2005 to help investigate the relationship between intended and actual length of stay. Results from new IPS questions are currently being researched.

Age groups and static populations

In England and Wales all age groups are treated in the same way except persons over 100 years old – who are processed as a single group. When publishing, however, all those over 90 are combined together due to reliability issues. However adjustments are made for special population groups that are not captured by migration estimates: members of the armed forces, prisoners and pupils in boarding schools. Other difficulties in measuring these sub-populations include them having multiple residences or residence periods within the estimation period of a year. These populations have specific age structures, which, as a whole, remain fairly constant over time. Therefore these groups are not aged on with the rest of the population. These groups are referred to as static populations. The static populations are more relevant when measuring internal migration and therefore are generally utilised when considering subnational population estimates. The armed forces are relevant at a national level because there are some armed forces living overseas and some foreign armed forces based in England and Wales.

In most countries the same methods are used for all age groups, as in England and Wales, but in New Zealand there is an interesting use of data for ages 0–9 years. Demographic estimates for people aged 0–9 at the national level are derived independently of the census from births, deaths and international migration data. These estimates are more accurate, at the national level, because of the virtually complete coverage of birth and death registrations, and international migration data. The accuracy of New Zealand census coverage at individual ages cannot be reliably estimated from post-enumeration surveys because of the limitations of sample size and associated sampling errors. Beyond nine years old the demographic estimates become increasingly susceptible to migration category jumping and therefore the census is considered a better estimate of the population aged ten and over. The ratio of demographic estimates to census-based estimates at the national level is then used to weight sub-national and ethnic population estimates at ages 0–9 years. In all cases the motivation for using a different method for different age groups is the availability of quality data at that age group. This method would not be appropriate for use in England and Wales because we do not have accurate enough information on international migration at individual ages, especially under ones.

In Australia estimates of the Aboriginal population are not calculated in the same way as the rest of the population. This is because the quality of data on births, deaths and migration for Aborigines is not of the same high quality as for the rest of the population. Therefore the census is used to create historical estimates of the indigenous population instead of the generally more accurate CCM.

Quality

Quality assurance of population estimates in England and Wales is assessed using demographic analysis such as sex ratios, birth/mortality rates and every ten years by comparison with the census.

Different countries use different methods to assess the quality of population estimates. The quality assessment in Portugal, as in England and Wales, is based on demographic analysis and comparisons against

results from the most recent census. In many countries estimates are produced from alternative sources to check against population estimates. In France and Greece these comparisons are made against administrative sources. In Estonia and Latvia they compare against different registers and in the USA, Australia, Ireland and Lithuania a comparison is made against the census. In France estimates are produced at the local level (departments and regions) using administrative records (electricity files, council tax) which are then used to check the quality of the local estimates derived from the census. ONS started annual quality assurance against administrative sources in their mid-2004 estimates. However, interpretation of the results is difficult and further research is ongoing as part of the IMPS project.

A census is a benchmark where the quality is very high and it is often used as a 'gold standard'. A large amount of work has been carried out by ONS in England and Wales to attribute the difference between the rolled forward estimates from 1991 and the census in 2001, particularly matching studies in Manchester and Westminster and a further 30 detailed local authority studies under the IMPS project

Summary

This section has compared the process used to produce population estimates in England and Wales with the processes used in other countries. There are some differences in methods used to compile population estimates, however these are mainly due to the availability of data sources. The main difference between countries is the availability of a population register. A population register is used to calculate population estimates in most countries that have one (Germany and Italy being exceptions). The advantage of population registers is the population coverage they can obtain. The CCM, as used in England and Wales, is highly dependent on the quality of the data used to measure each component. The data used to measure births and deaths in the UK is from civil registration and is of very high quality. However, the key difficulty in calculating population estimates with or without a population register is the measurement of migration.

The measurement of migration is difficult and each country will try to utilise the data source which can provide the highest quality migration estimates. Examples of data sources are population registration data, administrative data such as passenger card data or survey data such as the IPS. Combinations of these sources are also used. The definition of migrants also differs between countries. In England and Wales an international migrant is defined as someone who changes his or her country of usual residence for a period of at least a year, so that the country of destination becomes the country of usual residence. Twelve months is widely used because it complies with the UN definition but there are some exceptions including the USA and Italy. These examples occur when the migration data source does not use a 12-month definition. As a result it can be difficult to differentiate between short-term and long-term migration.

ONS and ABS make adjustments for migrant/visitor switchers. ABS has evidence from passenger cards, which is not available in England and Wales, that can provide accurate estimates. ONS may be able to make similar estimates when data from the e-Borders Programme are available. The same methods are used for all age groups in most countries. New Zealand is an exception and calculates estimates for people aged 0–9 years differently to the rest of the population. The availability of quality data at a particular age group will dictate the method used for that age group. The availability of alternative data sources will also dictate the quality assurance process. In many countries the benchmark for quality purposes is a census.

Recent improvements for calculating population estimates in other countries

Several countries have made changes to the methods of calculating population estimates recently. New Zealand moved from a *de facto* residence definition to usual residence for their census in 1996. This was to bring their statistics in-line with international practice.[12] In Australia a new definition has been proposed that classes a person as a usual resident if they have spent more than 12 months within Australia in a 16 month period. This new definition may help ABS measure international migration better by taking account of the number of people who migrate to Australia but return temporarily to visit family in their country of origin. These people may otherwise be classified as migrant switchers if they leave Australia before a year is up because they haven't been in the country for twelve months continuously. This approach may be useful in England and Wales if more information becomes available to measure international migration but is not possible at this time.

France has commenced a five-year rolling census. This rolling census samples a fifth of the population each year and therefore a total population estimate will not be available until the first five years have been completed. The methods for estimating migration will change with improvements to net migration estimation, particularly for in- and out-flows of students and EU-citizens which are not currently estimated accurately because these groups don't need to apply for a work permit. Data sources will include administrative data from the Ministry of Interior, information gathered in the census bulletins on the date of entry into France, the place of residence at the time of the 1999 census and income tax data. They would also like to use data from foreign censuses in order to estimate the number of French citizens living abroad. This may help in the estimation of emigration and ONS could study the results of any investigation to see if similar methods could be implemented in England and Wales.

There are a number of problems associated with retrieving information on foreign-born population estimates in foreign censuses. Information on flows is difficult to derive from data on stocks, there may be problems with the availability of data, problems in comparability due to definitional differences and difficulties in getting quality information on a particular subset (that is, French citizens) of the foreign born population. This final difficulty will be compounded for England and Wales as information may refer to people from the UK or Great Britain. Use of census data from overseas is being considered as part of the IMPS project.

The population estimates currently produced in France are provisional. In 2009 the mid-period estimate will be calculated using the average population estimate derived from the five most recent census surveys. that is, the population on 1 January 2006 will be available from the first five years of the census (2004–2008). This new estimate will replace the previous provisional estimate and will be considered the final estimate.

Hungary have been able to incorporate migration into their population estimates for the first time in 2001. In contrast in the same year Estonia decided that migration estimates were not of sufficient quality to be used in the population estimates and are no longer used. Again in 2001 the NSI in Latvia were given access to the administrative population register under the Ministry of the Interior which has enabled them to improve the population estimates.

Although countries have made changes to methods used to estimate population, most have been minor adjustments. The only reason for not using CCM to estimate the population is the presence of a population register. The only country to change from CCM to alternative method is France which has radically changed the mode of census collection to a five-year rolling census.

Other methods of estimation

There are a number of other methods beside CCM which can be used to estimate populations. This section describes the advantages and disadvantages of some of them.

The 'Housing Unit' method is the most commonly used method for estimating the population of small areas in the United States.[13] The population is estimated by multiplying the number of households by the average number of people per household and adding people living in communal establishments. If it were possible to use high quality data on numbers of households then the national population could be estimated very accurately but this is very rarely the case. ONS is currently investigating the use of housing stock in the measurement of population estimates in England and Wales.

As the exact number of households is not known, proxy information can be found from utility data (such as new customers to electricity companies) or data on construction such as planning applications. When construction data are used, a census acts as a base and the number of new households as estimated from the proxy data is added to the base and the number of demolitions subtracted. The number of households can be acquired by applying an occupancy rate to the estimate of housing stock. If utility data are available then the number of households can be estimated directly from the number of active residential customers.

The best source for the number of people per household is the census. Estimates of average household size between census could be calculated from large scale social surveys such as the Labour Force Survey (LFS). However, the LFS uses population estimates for grossing, potentially creating circularity, and response rates vary with household size so unweighted estimates may be biased. If possible, estimates of housing stock and number of people per household could be made by housing type and/or local area as the trends for different types will vary.

The 'Ratio Change' method is used in the Small Area Population Estimates published by ONS.[14] The CCM was evaluated together with two other short-listed methods – Apportionment and Ratio Change. The Apportionment method breaks down the population of a larger area into smaller areas using an indicator of population, for example from administrative sources. Inherent in this method is the assumption that the relationship between the indicator of population and the true population is the same for each small area within the larger area. While the evidence was not overwhelming, over a range of criteria, there was evidence that generally the Ratio Change method produced better small area population estimates than the CCM. The estimates produced using the Apportionment method were perceived to be the poorest. One of the main disadvantages with the CCM for producing small area estimates is the lack of quality international migration estimates at small area level.

The main assumption behind the 'Ratio Change' method is that, for each area, the data should have a consistent relationship with the true population over time. The ONS estimates were produced by applying change ratios to a ward estimate of the population base (the mid-2001 Census Area Statistics – CAS ward estimates) using a combination of administrative sources. Before applying these change ratios some static population counts are subtracted comprising armed forces and prisoners, and added again after these counts are constrained to the LA level mid-year estimates minus the special population. This method is highly dependent on the quality of the data used to calculate the change ratio. This method has subsequently been used to produce Super Output Area (SOA) estimates.

Regression methods for population estimation include the ratio-correlation method and the difference-correlation method. In these methods changes in population are related to changes in indicators

---

of population change such as school enrolment, tax returns, car registrations. Regression methods are based on the change in proportion of an area.

The Housing Unit method and the Ratio Change method are primarily used for producing estimates at small geographic levels but they can be adapted to national level estimates. These methods are generally applied when estimating small area populations because the data required for the cohort component methods is unavailable – in particular quality data for international migration are difficult to obtain for small areas. Conclusions drawn on the accuracy of methods for estimating small area populations cannot be assumed to hold for national population estimation.

Conclusions

In the absence of a population register the cohort component method is the best method for estimating the population. Although other methods of estimation are available they are used to estimate small areas because quality data for migration are not readily available or not of sufficient quality.

There are a small number of differences between the methods used by ONS to calculate population estimates in England and Wales and methods in other countries. The differences in method are:

- **The use of usual residence rather than de facto residence in population estimates**
  ONS is currently using a usually resident definition residence. This is the accepted international standard. If better information were to become available for measuring migration, for example through the e-Borders Programme, then investigation into the advantages and disadvantages of using a definition which takes into account long-term migrants leaving the country for a short period of time, such as that proposed by ABS may be suitable.
- **The use of census rather than a population register as a population base**
  Generally countries either use the most recent census or a population register as the population base. As England and Wales have has no population register it must use the census. Population estimation is likely to be more accurate in countries where the census is taken more frequently as the quality of the census as a benchmark reduces over time.
- **Different data sources are used by different NSIs particularly with respect to migration**
  NSIs will use the most appropriate data source at their disposal and these are not consistent across countries. It is not in the scope of this article to identify alternative data sources for use in the CCM. This confirms that ONS makes best use of the data sources available. If an alternative data source, such as migrant worker registration or e-Borders, were to become available it should be investigated.

- **Adjustments for visitor switchers are made in England and Wales and Australia only**
  The current methodology for migrant/visitor switchers is not based on empirical evidence. The only country that does make adjustments for migrant/visitor switchers has evidence from passenger cards which can provide accurate information on the scale of the problem. If data from passenger cards or the e-Borders Programme were made available it would be possible to make a more accurate adjustment for migrant/visitor switchers in England and Wales.
- **Some countries utilise different data sources at different age groups**
  In most countries the same methods are used for all age groups, as in England and Wales, but in New Zealand demographic estimates for people aged 0–9 years at the national level are derived independently of the census from births, deaths and international migration data. The ratio of demographic estimates to census-based estimates at the national level is then used to weight sub-national and ethnic population estimates at ages 0–9 years. Where different methods or data sources are used for different age groups the motivation is the availability of quality data at that age group.
- **From the information collected static populations are only utilised in England and Wales**
  Static populations, in general, are more of a problem when measuring internal migration and therefore are used more when considering subnational population estimates. The armed forces are relevant at a national level in England and Wales and they are potentially relevant in Germany which has a large population of foreign armed forces. No information has been obtained from Germany and therefore this is an area that could be followed up in more detail.

Key findings

- In the absence of a population register, the CCM is the best method for estimating the national population.
- The driver of population change in England and Wales has changed in recent years. Change is now driven by migration, a component which is increasingly difficult to measure.
- The base population, from which annual estimates are rolled forward, is normally a census (as for England and Wales) or a population register.
- Data sources used in the application of the CCM are different around the world, in particular in the estimation of migration.
- Adjustments for Migrant and Visitor Switchers are only made in England and Wales and Australia.
- Research undertaken in Australia proposes using a usual residence rule of 12 months out of the last 16. This is of interest for developing ONS's population definitions.
- Unlike in England and Wales, some countries – most notably New Zealand – use different data sources for the estimation of different age groups.
- Estimates of 'static' populations are only made in England and Wales.

Acknowledgements

Note

EU-15 is a term used to indicate the member states prior to May 2004 when ten new member states acceded the EU

References

1. Jefferies and Fulton (2005) *Making a Population Estimate in England and Wales.* Available at: www.statistics.gov.uk/about/data/methodology/specific/population/PEMethodology/
2. Australian Bureau of Statistics (1995) *Demographic Estimates and Projections: Concepts, Sources and Methods*. Available at: www.abs.gov.au/asstats/ABS@nsf/66f306f503e529a5ca25697e0017661f/57917ee0c347fb50ca25697e0018fb50!OpenDocument
3. Horsfield G (2005) International Migration in (R Chappell ed.) *Focus on People and Migration,* Palgrave Macmillan: Basingstoke, Chapter 7, pp 115–119. Available at: www.statistics.gov.uk/focuson/Migration
4. Jones J and Chappell R (2004) European wide issues in population statistics. *Population Trends* **118**, 17–22.
5. Smith C W and Jeffries J (2006) Population bases and statistical provision: towards a more flexible future? *Population Trends* **124**, 18–25.
6. Bas de Vet (2003) Final summary report concerning the methodology of regional population estimates, A paper to the Voorgurg Regional Population Workshop
7. Poulain M, Perrin N and Singleton A (2006) THESIM, Towards Harmonised European Statistics on International Migration, UCL Presses Universitaires de Louvain [ISBN: 2-930344-95-4].
8. Eurostat (2003) Demographic Statistics: Definitions and methods of collection in 31 European countries, Luxembourg, Office for Official Publications of the European Communities, p 23. [ISBN 92-894-6051-2] available at www.uni-mannheim.de/edz/pdf/eurostat/03/KS-CC-03-005-EN-N-EN.pdf
9. United Nations Statistics Division (1998) *Recommendations on Statistics of International Migration*, Series M No.58, Rev.1. New York.
10. Schengen Convention. Available at: http://ec.europa.eu/justice_home/fsj/freetravel/frontiers/fsj_freetravel_schengen_en.htm#
11. Statistics New Zealand, (2002) *Information about the population estimates*. Available at: www2.stats.govt.nz/domino/external/omni/omni.nsf/outputs/Population+Estimates
12. Ryan M (1999) *The effects of changing from de facto population estimates to* resident population estimates. Demography Division of Statistics New Zealand. Available at: www.stats.govt.nz/products-and-services/Articles/pop-est-changes-Mar99.htm
13. Smith S K (1986) A review and evaluation of the housing unit method of population estimates. *Journal of the American Statistical Association* **81** (394), 287–296.
14. Small Area Population Estimates. Available at: www.statistics.gov.uk/about/methodology_by_theme/sape/default.asp
15. E-borders Programme. Available at: www.homeoffice.gov.uk/documents/ria-data-capture-200705?version=1