A cross country review of the validation and/or adjustment of census data

Rebecca Newell and Steve Smallwood

Office for National Statistics

Abstract

This article reviews existing procedures employed by various countries in the evaluation of, and/or adjustment, either of census data, or of population estimates based upon census data. The work was carried out to ensure all potential demographic techniques are considered by the ONS for the post census evaluation process of the 2011 Census.

Contents

A cross country review of the validation and/or adjustment of census data	115
Abstract.....	115
Introduction	117
Approaches used in demographic analysis	118
Non-demographic methods of direct census adjustments	121
Non-demographic methods of post census adjustments for population estimates	122
Use of additional sources in census validation	124
Alternative types of estimation	125
Summary and conclusions	126
References.....	127

Introduction

The purpose of this article is to review existing procedures employed by various countries in the evaluation of, and/or adjustment either of census data or population estimates based upon Census data. The work was carried out to ensure all potential techniques are considered by the ONS for the post census evaluation process of the 2011 Census.

The focus of the work was to consider in particular what demographic techniques were used when considering census results, and whether changes were made either directly to the census database, or as subsequent additions to aggregate results, for the purposes of estimating populations by age and sex. Several procedures have been developed by individual countries over the past few decades and play an important role in understanding and evaluating the coverage of census data. One of the major problems, however, is that it is very difficult to know when to make an adjustment and what the level of adjustment should be. If you had that information you would be likely to have high quality information about the age and sex of the population and it would probably not be necessary to take a census.

There are a number of different methods used to analyse and adjust census data. Some of these can clearly be classed as statistical, such as using 'capture/recapture' techniques or direct macro or micro comparisons with other data sources, while some are more demographic. Demographic analyses are methods that allow us to measure the dimensions and dynamics of populations such as size, composition and spatial dynamics, and how these features change over time. Such indicators are often derived from administrative records of events like births, deaths, marriages, divorces, diseases and employment. The underlying principle behind using demographic methods to validate census data is that demographic analysis can provide measures, which are independent of the census population counts. This can be used as evidence of how many people were missed on census day and may provide a way of making appropriate revisions to population counts where major discrepancies are found.

The use of demographic analysis in census validation is commonly cited in the literature as a 'possible alternative' or 'independent measure' of population counts at the national level (macro level analysis), but most papers do not provide any further information on the actual methods postulated. A review by Don Kerr¹ during the 1990s presented the various procedures used in estimating net undercount of censuses (prior to the millennium round of censuses) in Canada, the USA, Britain and Australia. Data collection methods were briefly described; micro-level procedures (record linkage) outlined; and macro-level procedures (demographic analysis), examined and found to vary in importance and application between countries. Relevant methodologies are discussed below and under the section headed 'Post Census Adjustment Using Demographic Methods'. However, since Kerr's review, little research on the explicit use of demographic techniques for evaluating census data is currently available, and even more sporadic are published examples of the application of demographic analysis in validating or adjusting census figures. From the published literature, there are a number of inter-related techniques that predominantly focus on constructing population cohorts.

Approaches used in demographic analysis

Cohort style approach

A cohort style approach is where estimations are made of expected numbers of the population from known information on births and mortality rates (and in some cases known migration). This works on the simple principle that a known population aged x will be aged $x+t$, in t years. These analyses can be used to examine internal consistency of population counts in terms of their conformity with expected age, period and cohort patterns of coverage; examples are summarised below.

Charlton *et al.* (1997)² stated that, 'cohort analysis is a way of providing estimated population counts, independent of all censuses'. They traced cohorts based on Government Actuary life tables at the national level (England and Wales) between Censuses 1971, 1981 and 1991. Results from the cohort analysis provided estimated population counts most similar to the official population counts for 1981, and concluded that 1981 provided the best base from which to roll forward estimated population counts for the 2001 Census benchmark. The benefits of utilising this type of comparative methodology were further illustrated within their analysis, which illustrated the problem of rolling forward the undercount of the 1960s inflow of migrants, where more people were found in the census than the cohort analyses predictions for ages 30-49 in 1971, 40-59 in 1981, and 50-69 in 1991. This approach also clearly demonstrated the emerging phenomenon of 'missing young men' from Census 1981 onwards.

Takami (2003)³ addressed a similar issue of the 228,561 'age-unknown' persons to appear in the Japanese 2000 Census (i.e. those who were counted in the census but 'age' was missing). An age restricted cohort analysis was used to compare the results of the 2000 Census and the previous 1995 Census. The cohort consisted of persons under 50 years of age, as the number of deaths was believed non-negligible in five year age groups beyond 50. The analysis showed discrepancies between the expected reduction in age-specific populations due to deaths and moves abroad, and the numbers actually missing for the 2000 Census age groups. They concluded that 'age-unknown' persons were most likely to be concentrated around the 20-29 year age group, as the number missing from this age group in the 2000 Census compared to the expected reduction due to deaths and moves abroad, was greater than for all other age groups. Although evidence was found of age-related discrepancies, no formal adjustments were made.

An example of more recent work relates to the quality assurance project carried out by ONS for the 2001 One Number Census (ONC)⁴. This involved comparison of ONC counts against diagnostic ranges derived from rolled-forward population estimates (as advised by Charlton *et al.*, 1997²) and the best available aggregated administrative data sources (such as birth registration and pensions data). Where differences between ONC estimated counts and diagnostic ranges existed, extensive checks of the ONC results and diagnostics were undertaken. However, as the authors stated, 'it was never the intention for the census to be adjusted to these diagnostics' and as discussed later, a decision subsequently questioned by academics (see 'Use of Additional Sources').

Demographic age-sex ratios

A second body of inter-related literature refers to comparisons with demographic age-sex ratios. These are predominantly derived from the cohort analysis described above. These provide auxiliary evidence for the growing phenomenon of 'missing men'.

Sex ratios will be affected by *relative* undercount, or overcount, of males and females. One particular piece of analysis that can be carried out is to construct from births data and mortality rates the numbers expected to be alive. The overall result of these calculations allows age/sex ratios to be calculated in the absence of migration. For example, work presented at the European Association for Population Studies (EAPS) conference in Barcelona⁵ demonstrated for England and Wales a dipping of the male to female sex ratio below 100 from age 20 onwards, as opposed to the natural reversal in the sex ratio from approximately age 50 onwards. **Figure 1** shows the UK Census 2001 age sex ratio against a synthetic population from historical cohort data. The synthetic population was derived from the population of England and Wales born estimated to be alive today, and created by applying historical age-specific mortality rates from Government Actuary Department (GAD) to births.

Further the population figures created can then be compared with the numbers found alive in the Census who were born in the country. This number should always be lower than the estimate of those expected to be alive. It can be inferred that the difference are those that are alive outside the country. The sex ratio for this group can also be computed. However, this latter calculation is not completely independent of the Census as Census data is used to infer the sex ratio of the England and Wales born population abroad. Rather it raises the issue of the plausibility of the resulting ratios.

General findings from around the world relating to this type of analysis⁶ elicit similar patterns between census data and demographic age-sex patterns to those depicted in Figure 1. When data is further disaggregated, by ethnicity for example^{7 8} the discrepancy between sex ratios at younger ages is even more evident.

Figure 1 Sex ratio patterns for England and Wales in 2001 from different estimation methods

Further examples of the use of demographic analysis

The US Census Bureau have used demographic analysis since 1960 to assess coverage levels of the census by providing estimated population counts, independent of census population counts by age, sex and ethnicity. These demographic estimates are derived from a culmination of births, deaths, legal immigration and emigration figures, supplemented by data on armed forces, and estimates of undocumented immigration. However, due to the lack of completeness of birth registration data prior to 1940 in the USA, administrative Medicare enrolments are used to estimate the population aged 65 and over.

Wolter (1990)⁹ and Bell (1993)¹⁰ explored ways to combine results from US demographic analysis with estimates of census undercount (derived from Dual System Estimation techniques, which are discussed later in this article). Specifically, alternative strategies of incorporating information on sex ratios were investigated, but the findings were complex and were not adopted by the US Census Bureau for the 2000 Census. This is covered in more detail under the 'Uncertainty Models' section below.

Post Census 2000, Robinson *et al.* (2002)⁷ provided a detailed analysis of the differences between the US Census 2000 counts and initial demographic analysis by age and sex. The results showed a clear undercount in the census figures, but that the under-coverage was disproportionate by age and sex, with higher numbers in the demographic estimates for both males and females in the age group 18-29. The findings were put forward as a discussion point for future assessments of the quality of demographic analysis estimates within the US.

Bycroft (2006)¹¹ explored data quality issues with New Zealand census population counts and those derived from demographic analysis, using sex ratios as their point of comparison. Census counts in New Zealand have shown steady decline in sex ratios since the 1960s, but this decline is not reflected in demographic population counts derived from changes in births, deaths and migration, raising questions over the reliability of more recent census disaggregation by age and sex.

Of the research papers and statistical institutions that indicated discrepancies between demographic analysis and census data, none have led to any formal adjustments to national census figures based on the findings from demographic analysis.

Post census adjustment using demographic methods

The majority of census taking countries in the developed world produce at least two sets of figures: initial census counts that are released soon after or within two years of the census, and some form of 'estimated resident population count', which is separate from the census counts. It is within these estimated population figures that certain countries have adopted some degree of demographic adjustment. However, information on this is limited but key points are highlighted below.

The Australian Bureau of Statistics (ABS) uses 'demographic adjustments' at the sub-national level (Australian States) by age and sex within the preliminary process of producing their 'initial' Estimated Resident Population count (ERP). The demographic adjustments are derived from data on demographic events compiled since 1921 (births, death and overseas migrations), using conventional cohort component analysis, resulting in estimates for the population aged 0-69.

A further comparison is made with the sex ratio in each age group of people enrolled for Medicare, which they believe to be especially useful at the very youngest and older ages where accuracy of the Medicare data is deemed most reliable.

A combination of weighted results for the Post Enumeration Survey (PES) and Medicare data are used to adjust the age groups 70 years and over, whereas a weighted comparison of all three methods (PES, demographic analysis and Medicare data) are used for the population aged 0-69. Unfortunately, the criteria or confidence intervals upon which they base their decision to make direct changes are not stated in the technical papers provided^{12 13}.

Australia later produced 'final' ERPs that included the addition of Residents Temporarily Overseas (RTOs) on census night (discussed in more detail later). Further sub-state demographic adjustments, which include checks against Electoral Roll data and armed forces adjustments, plus or minus the standard birth, death and migration figures since the elapse of the census date are administered. Essentially ERP undergoes two stages of adjustment, resulting in an 'initial' ERP and a 'final' ERP. The main difference between the two is the amount of information used in each adjustment with the belief that as time goes on, more information becomes available for subsequent adjustments and hence the best possible estimates are produced.

In the production of Population Estimates for Statistics Canada, the initial estimates of net undercount (derived from the Coverage Error Measurement Program), are formally adjusted where implausible discrepancies are found when compared to demographic components of population change (births, deaths, migration and immigration)¹⁴. An example of where a direct adjustment was made relates to the 1996 Census, where the Reverse Record Check (RRC) results for children aged 0-4 provided an estimated net undercount significantly higher for female children with no obvious explanation. Consequently, the RRC and census results combined gave an implausible figure and subsequently, the number of females was revised to correspond with the demographic age sex ratio¹.

In a similar fashion to Canada, Statistics New Zealand adjusts their 'estimates of net undercount' for ages where large discrepancies exist. For example, adjustments were made for ages 0-9 where substantial discrepancies were found when comparing to the standard demographic components of population change¹¹, but again, no criteria for the adjustment is cited.

Non-demographic methods of direct census adjustments

In 2001 the United Kingdom developed the One Number Census, which included direct adjustments to raw census count figures, on the basis of Dual System estimation (described below), designed to result in only one set of census population figures being released; the first of its kind¹⁵. At present the only other country, apart from the United Kingdom, to make any form of 'direct' adjustments to their raw census count figures is Canada, with an overview of the methodology presented below.

Direct adjustments to the census database are administered by Statistics Canada using their post-Census 'Dwelling Classification Study' (DCS), formally known as the 'Vacancy Check'. There are actually four coverage studies used by Statistics Canada (DCS; Reverse Record Check (RRC);

Automated Match Study (AMS); and Collective Dwelling Survey (CDS) but the DSC is the only one that results in the addition of households and persons to the census database.

The DCS adjusts for unoccupied and non-response dwellings (for example classification errors). If a dwelling was occupied in the DSC, one of two adjustments is made to the census database:

1. If the dwelling was listed as 'un-occupied' in the census, then persons deemed missing are added to the census database using the random addition technique (a form of imputation).
2. For non-response dwellings (for example, occupied dwellings for which a completed census was not received), an adjustment is made by creating a new household size for them on the census database.

The results of the remaining three coverage surveys (RRC, AMS and CDS) relate to the calculation of the base population for the post census population estimates as opposed to direct census adjustments¹⁴.

Non-demographic methods of post census adjustments for population estimates

Post Enumeration Survey (PES)

Measured levels of census undercount has increased as a whole over the past few decades, with priority and focus being given to measuring this differential undercount. Consequently, most census taking developed countries now undertake some form of coverage assessment and adjustment, usually in the form of a post-enumeration survey (PES). Detailed examples from each country (Australia¹⁶, Canada¹⁷, Japan¹⁸, New Zealand¹⁹, United Kingdom²⁰, and USA²¹) can be found in respective NSI technical documents.

Coverage Surveys

The Reverse Record Check is a comprehensive record-linkage method used in Canada by which official gross under and over-coverage is estimated. The system that processes the addresses of the selected persons is automated; however, it is not clear from the technical paper what source these 'addresses' come from. After the sample is taken, persons who should have been enumerated and were not, are then traced and interviewed and the results used to determine the number and characteristics of unenumerated persons. Some limitations to this methodology exist where non-sampling error cannot be quantified in the estimates (that is, those who were untraceable and could not be interviewed). Coverage surveys are also used to identify over-coverage, where people can be counted more than once. The Automated Match Survey (AMS) detects erroneous enumerations by a combination of matching procedures, to identify duplicate households in the census database (for example, persons enumerated in more than one household within the same region). The Collective Dwelling Survey (CDS) estimates over-coverage resulting from persons enumerated as usually resident in non-institutional collective dwellings who are also enumerated at private dwellings. The Reverse Record Check (RRC) was designed to measure over-coverage from all sources (RRC, AMS and CDS). If over-coverage is detected in the AMS or the CDS, then person records are removed from the RRC to eliminate multiple counting of over-coverage between the three coverage surveys. However, from the 2001

Census onwards the CDS component was dropped, and over-coverage estimates for this group (historically very small) were made instead as part of the RRC¹⁴.

Dual System Estimation (DSE)

Dual system estimation is another method that can be used to generate population estimates from census data. The USA ruled out the possibility of using a record-linkage study like the Reverse Record Check (RRC), concluding that such procedures were inappropriate in a country that carries out a decennial census, as tracing records from ten years previously would be too problematic. Instead, the USA use a technique called 'Dual-System Estimation' (DSE) where the PES consists of two basic samples: the P sample (area sampling) and the E sample (sample of census enumerations from census database). The proportion of the P sample actually enumerated in the PES provides an estimate of gross under-coverage and is validated through re-listing, re-interviewing, follow-up interviews and record matching. The estimates of under and over-count from P and E samples are then combined to provide a 'dual-system estimate' of net undercount. More detailed information of the procedure can be found in Kerr (1998)¹. However, DSE is subject to both sampling and non-sampling errors, although continuous efforts are being made to minimise these errors. Alternatives to the dual-system estimates were tested prior to the 2000 Census, in the development of an Integrated Measurement Program, but the attempts were unsuccessful^{1 22}. As we have already seen, this methodology can also be used to directly adjust the Census database, for example, in the United Kingdom's 2001 Census¹⁵.

Residents Temporarily Overseas (RTOs)

Both Australia and New Zealand utilise international migration data derived from passenger arrival and departure cards in the twelve months following census. New Zealand's migration data is further supplemented with the capture of electronic passport data during the same 12 month period. In both cases, international migration data is analysed for duration of stay overseas to determine whether persons were missing on census night. Adjustments are then made to Estimated Residential Population counts (EPRs) accordingly^{13 23 24}.

Due to a change in classification for net overseas migrations (NOM), now defined as being in/out of Australia for 12 out of 16 months, a new two-part method has been proposed for RTO adjustments of future censuses in Australia²³. At present, New Zealand have not changed their NOM definition but are closely following the outcome of ABS's change in definition and subsequent data output.

Explorative work in New Zealand¹¹ investigated a new method to improve measures in the change in population due to external migration. At present, no adjustments are made on this basis, but could be incorporated in future censuses. As Castle and Miller (2003)²⁵ point out:

'... the experience of migration and of living in another country often leads to modification of the original plans, so that migrants' intentions at the time of departure are poor predictors of actual behaviour'

Bycroft's aim was to use total net migration over a period of time with corrections for incomplete short-term travellers at the start and end of the period, thus adjusting for any 'change' in travel intentions as opposed to taking peoples 'intentions' as definitive. Preliminary results suggest the real contribution of migration to population change is not measured well by current Permanent and

Long Term (PLT) migrant classifications and the new method provides a promising avenue for further research into the effects of misclassification of short and long term migrants.

Use of additional sources in census validation

The United Kingdom

The use of additional data sources in the estimation process has been continually suggested within the literature, but as yet little has been taken forward. Prior to the One Number Census 2001, Chappell *et al.* (1997)²⁶ stated that administrative sources for the population as a whole and for certain population subgroups needed to be identified along with their quality and availability. Following this, a variety of demographic and administrative record checks (Child benefit; Pension data; MYEs; Home Office prison data; HESA and LSC data) were used to compare the results of Census 2001. No demographic adjustments were made but two of the administrative record checks led to direct adjustments in the ONC process for students' halls of residence figures and DASA data for the armed forces⁴. In summary, Diamond (2003)⁴ stated that child benefit and pension data at the time of the 2001 Census were deemed the most accurate and nationally consistent administrative sources available in the UK.

A review of the ONS One Number Census count and quality assurance checks was published in 2007²⁷. The collection of independent demographic data and administrative records used with the intention of diagnosing where the census estimates had fallen considerably outside what was expected were discussed. The author questioned why, when over 20 per cent of the 14,000 comparisons made fell outside the more lenient 'demographic and administrative diagnostic range' compared to the 95 per cent confidence intervals for the ONC design, subsequent adjustments were not made to the estimates. The recommendation put forth was that a third source from administrative records *should* be used in future adjustments where discrepancies outside tolerance levels exist. However, no third administrative data source with full population coverage that is more reliable than the census exists.

Looking ahead to the 2011 Census, while quality issues with administrative sources exist, ONS have postulated that the use of additional sources is vital for the 2011 Census and have highlighted some potential sources for use in the estimation process and diagnostic ranges: 'Visitor Data'; the 'Census Household Frame'; and, the use of improved survey data, such as the 'Integrated Household Survey' (IHS) which should be available for 2011²⁸. In addition, for the first time in 2011, information on birth and death rates by area will be used to help quality assure census data.

Outside the UK

Both Australia (ABS) and the United States use Medicare data to check the coverage in their population aged over 65 and to make subsequent adjustments to their population estimates where necessary^{7 10 12 29}. The ABS also looks at their Electoral Commission (AEC) population distribution against their census population distribution and make further adjustments at the subnational level where clear deviances exist¹¹.

Japan has investigated Basic Resident Registers (BRR) as a validation check, which cover only those who have Japanese nationality and compare the population census and the BRR with

respect to persons of Japanese nationality. At present no adjustments are made due to questions over the reliability of the BBR data (i.e. non-reporting of moves)³.

Although Sweden uses a population register as opposed to a traditional census, they have looked at using mortality data to estimate indirectly over-coverage by studying the differences in mortality between immigrants and the native population. The mortality levels for immigrants are compared with those of people born in Sweden after standardising for various background variables. When immigrant mortality in the population registers is very low compared with native Swedish mortality, relative over-coverage in the Resident Total Population (RTP) is assumed. But as the author states, this method has its inadequacies since factors other than errors in the registers may lie behind the differences³⁰.

Alternative types of estimation

Triple System Estimation

Triple system estimation has been suggested as a method to counterbalance the standard problem of dependence in dual system estimation, where a third list is matched to both Census and PES survey data^{27 31}. In both the 1990 and 2000 US censuses, expected sex-ratios were used as the third source in the estimation process, but the data were not formally adjusted in published Census counts¹⁰. In the UK, it was concluded³¹ for the 2001 Census that although some properties are missing, a combined PAF (Royal Mail Postcode Address File) and Census list is currently the best source available, but realistically the resolution is not sufficient enough to act as a third data source at the individual level.

Uncertainty Models

Research during the 1990s explored ways to combine results from demographic analysis with dual-system estimates of net undercount obtained from the PES^{9 10}. The original work by Wolter (1990)⁹ used sex-ratios to introduce systematic revisions of dual system estimates. Bell (1993)¹⁰ then extended this work by attempting to modify the 1990 PES dual-system estimates using demographic estimates of national sex-ratios by age. The U.S. Census Bureau believes sex ratios to be a robust tool for Census evaluation, and although raw figures are not actually adjusted, post Census figures incorporate these methods¹.

Robinson *et al.* (1993)⁷ then developed 'uncertainty models of demographic analysis' that evaluated the uncertainty associated with the measurement of each component of the demographic estimates, including assigning confidence intervals to estimated sex-ratios by age. Their work was subsequently incorporated in the development of US demographic estimates at the subnational level for Census 2000.

Bayesian Methods

The application of Bayesian methods to address uncertainty and 'missingness' in census data, and the framework in which it 'borrows strengths' from additional sources (combining data) is positive in terms of its statistical theory. Bayesian methodology follows three basic principles:

1. Specification of a probability model that includes some prior knowledge about the unknown parameter values (for example, missing persons)

2. Updating of knowledge about unknown parameters by conditioning this probability model on observed data (for example, demographic measures)
3. Evaluate the fit of the model to the data and the sensitivity of the conclusions to the assumptions.

The few relevant publications that have attempted to explore the use of Bayesian methods in post census population estimates are:

Dick and You (1997a; 1997b)^{32 33} attempted to create a more flexible model compared to conventional population estimates where undercoverage rates compared to the national rate in Canada were fixed across provinces. Instead, their model permitted different levels of underlying provincial under-coverage to exist, by allowing the exchangeability of provinces to be examined under a hierarchical Bayesian model³². They concluded that some improvement over direct population estimates was possible, but that the methodology was exceeding technical for little, if any quantifiable gain.

Elliot and Little (2000; 2005)^{34 35} evaluated a number of models that have been proposed within the USA to deal with post census adjustments, and chose what they believed to be the most robust and placed it within a Bayesian framework. Their Bayesian approach removed the problem of negative cell counts inherent in maximum likelihood methods (MLE), but provided no conclusive outcome as to whether their methods were any better than those already used. The direct benefits of using a Bayesian approach in census validation and or adjustment is currently inconclusive.

Summary and conclusions

Few countries have explicitly used demographic analysis in their evaluation process, and very few use such analysis to adjust either Census counts or population estimates based on the Census. Those which have either used the findings to highlight where possible discrepancies lay within their published census counts (but do not make direct adjustments), or else utilised the demographic discrepancies from the census counts to make post census adjustments for various subpopulations at the national level in their 'Estimated Population Count' procedure.

There are known limitations to the use of demographic analysis: in fact, if demographic analysis was able to give a completely reliable answer to population numbers by age sex and geography, then a census would not be necessary. Aggregate administrative data may contain errors or be incomplete; thus linkage to such data, assuming legal gateways are open, is reliant on the quality of the linkage and interpretation of the unlinked cases. Survey data may not be representative of population characteristics due to non-response, but even if it is representative, will have confidence intervals wider than the accuracy required for Census at low levels of geography. Further, weighted survey data will be weighted to population estimates – the very estimates for which the census is needed. Rates derived from vital events give an indication of potential issues with census populations, but not a precise level for adjustment.

However, the importance of demographic analysis should not be overlooked. Sex-ratios are a robust tool for evaluating age-sex patterns found within the census and should be a consideration in any post census adjustment. One of the major problems relates to confidence intervals around census data, both in terms what level of accuracy is required and whether adjustments actually

improve the accuracy, that is, when does one *really* need to adjust? Unfortunately, limited information was available on this issue from the National Statistical offices discussed.

The usefulness of additional sources of administrative data should also be considered, and utilised where obvious discrepancies are found. Ideally, one would use a third 'total population' data source to both cross reference and help validate census data, but as that 'elusive' third dataset does not exist within the UK a combination of smaller scale comparisons of subpopulation data could be used in an amalgamation of 'small but significant' adjustments where the census data falls short of expected or known profiles. In particular, this must be the evaluation of subgroups of the population where specific data sources for these subpopulations are in most cases fairly accurate (for example, the elderly, the young, school children, students and the armed forces).

References

- 1 Kerr, D (1998) 'A review of procedures for estimating the net undercount of censuses in Canada, the United States, Britain and Australia'. Statistics Canada, Research Paper, Demographic Documents; Catalogue no. 91F0015MIE-No.005.
- 2 Charlton, J, Chappell, R and Diamond, I (1997) 'Demographic analysis in support of a one number census'. ONC ONS(SC)97/15, ONS, Census Division.
- 3 Takami, A (2003) 'Evaluation of the accuracy of the 2000 population census of Japan'. 21st Population Census Conference. Japan, Kyoto.
- 4 Diamond, I (2003) 'Key Issues in the quality assurance of the one number census'. *Population Trends* 113: 11-19.
- 5 Smallwood, S and De Broe, S (2008) 'Unravelling sex ratios, what can they tell us about the Census population estimates and recent trends in migration?' Paper presented at the EAPS conference in Barcelona 10-12 July 2008.
- 6 Smallwood, S and De Broe, S (2009) 'Sex ratios patterns in population estimates'. *Population Trends* 137: 41-50.
- 7 Robinson, JG, Ahmed, B, Das Gupta, P and Woodrow, KA (1993) 'Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis'. *Journal of the American Statistical Association* 88 (423) 1061-1071.
- 8 Robinson, JG, West, KK and Adlakha, A (2002) 'Coverage of the population in census 2000: results from demographic analysis'. *Population research and policy review* 21 (1-2) 19-38.
- 9 Wolter, KM (1990) 'Capture-Recapture estimation in the presence of a known sex-ratio'. *Biometrics* 46 (1) 157-162.
- 10 Bell, WR (1993) 'Using Information from Demographic Analysis in Post-Enumeration Survey (PES) Estimation'. *Journal of the American Statistical Association* 88 (423) 1106-1118.

- 11 Bycroft, C (2006) 'Challenges in Estimating Populations'. *New Zealand Population Review* 32 (2) 21-47.
- 12 ABS (2006) 'Census Undercount. Demographic Estimates and Projections: Concepts, Sources and Methods'. August 1999. Available at: www.abs.gov.au/ausstats/abs@.nsf/7d12b0f6763c78caca257061001cc588/F66C2B24DC1AE3FFCA25697E0018FC5A?opendocument
- 13 ABS (2006) 'Rebasing Australia's population estimates using the 2006 census of population and housing'. *Australian Demographic Statistics*, September 2006.
- 14 Statistics Canada (2004) 'Coverage: 2001 census technical report'. Statistics Canada Catalogue number 92-394-XIE. Available at: www.statcan.ca/english/census01/Products/Reference/tech_rep/coverage/offline%20documents/92-394-XIE.pdf
- 15 Office for National Statistics (2005) '*Census 2001 review and evaluation: One Number Census Evaluation Report*'. ONS, London.
- 16 www.abs.gov.au/websitedbs/D3310114.nsf/Home/census?opendocument#frombanner=GT
- 17 www.statcan.ca/english/census01/Products/Reference/tech_rep/coverage/offline%20documents/92-394-XIE.pdf
- 18 More detail available at: www.stat.go.jp/english/data/kokusei/index.htm
- 19 Available at: www.stats.govt.nz/analytical-reports/post-enumeration-survey-2001/chapter-three-post-enumeration-results.htm
- 20 ONS 2001 Census info available at: www.statistics.gov.uk/census2001/pdfs/oncguide.pdf
- 21 Available at: www.census.gov/dmd/www/refroom.html
- 22 Mulry, MH and Griffiths, R (1996) Comparison of Census Plus and Dual System Estimation in the 1995 Census test. AMSTAT. Available at: www.amstat.org/Sections/Srms/Proceedings/papers/1996_146.pdf
- 23 ABS (2009) ABS 2006 Census Rebasing Papers: Australian & State/Territory RTO's: Census Night 2006. *Internal paper; unpublished*.
- 24 Statistics New Zealand (2009) Post enumeration survey results. Available at: www.stats.govt.nz/analytical-reports/post-enumeration-survey-2001/chapter-three-post-enumeration-results.htm
- 25 Castles, S and Miller, MJ (2003) *The Age of Migration*. Palgrave MacMillan, Guilford.
- 26 Chappell, R, Teague, A and Diamond, I (1997) *Methodological research for a 'one number census'*. Office for National Statistics, Census Division. Available at: www.ons.gov.uk/census/2011-census/consultations/user-adv-groups/census-adv-groups/statistical-dev/meth-res-1-num-cen.pdf
- 27 Simpson, L (2007) 'Fixing the population: from census to population estimate'. *Environment and Planning A* (39) 1045-1057.

-
- 28 Abbott, O (2007) '2011 census coverage assessment and adjustment strategy'. *Population Trends* 127: 7-14.
 - 29 ABS (2009) ABS 2006 Census Rebasing Papers: Estimating Statistical Local Area (SLA) and Indigenous Estimated Resident Population (ERP). *Internal paper; unpublished*.
 - 30 Qvist, J (1999) 'Problems of coverage in the register of total population (RTB): Estimation of overcoverage by an indirect method'. *Statistics Sweden*, R&D Report 1999:1.
 - 31 Brown, J, Abbott, O and Diamond, I (2006) 'Dependence in the 2001 one-number census project'. *Journal of the Royal Statistical Society* 169: 883-902.
 - 32 Dick, P and You, Y (1997a) 'Bayes and census undercoverage'. Statistical Society of Canada, Annual Conference; Proceeding of the Survey Methods Section, Fredericton, New Brunswick.
 - 33 Dick, P and You, Y (1997b) 'A hierarchical Bayes analysis of census undercoverage'. Proceedings of the Statistics Canada Symposium 199; Work Session on Advances in Imputation, Canada: Ottawa.
 - 34 Elliot, MR and Little, RJA (2000) 'A Bayesian approach to combining information from a Census, a coverage measurement survey and Demographic analysis'. *Journal of the American Statistical Association* 95 (450) 351-362.
 - 35 Elliot, MR and Little, RJA (2005) 'A Bayesian approach to 2000 census evaluation using ACE survey data and demographic analysis'. *Journal of the American Statistical Association* 100 (470) 380-388.