



Development of Sample Surveys in Australia and New Zealand over the Last 50 Years

Dennis Trewin

Australian Statistician (2000-2007) and IASS President (1995-1997)
dennistrewin7@gmail.com

Abstract

This note outlines the development of sample surveys in official statistics in Australia and New Zealand since the inception of the IASS. It highlights the increasing need to not just rely on sample surveys although they remain important as a component of mixed mode methods. Accordingly, sample survey methodologists need to broaden their skills to maintain their relevance.

Keywords: sample surveys, mixed mode, non-sampling errors.

1 Introduction

Both countries have centralised statistical systems and the majority of sample surveys are conducted by the national statistical offices. Hence, this paper concentrates on developments in these two offices. Given the context of this issue of *The Survey Statistician*, the paper is also more about past history than recent history. A good reference for Australia is ABS (2005).

As an inaugural member of the IASS in 1973, I have seen many changes in survey methods over the last 50 years. Nevertheless, both Australia and New Zealand (NZ) were relatively mature in their adaption of probability-based sample surveys by 1973 under the leadership of Ken Foreman at the Australian Bureau of Statistics (ABS) and Steve Kuzmicich at Statistics New Zealand (SNZ). The importance of their work was recognised because, for much of their careers, they held quite senior positions and were regarded as part of the executive team as well as being chief methodologists.

By 1973, a national household survey was in place in Australia and had been held since 1960. It was used to conduct the quarterly labour force survey (LFS) and a series of supplementary surveys. In New Zealand, the LFS came later whilst they relied on registered unemployment data. In both countries, a number of special household surveys such as household income and expenditure surveys (HIES) had already been conducted. Multistage area frameworks were used with a mesh block as the primary sampling unit (PSU) in NZ whilst Australia used the larger Census Collectors Districts (CDs) and used field work to create blocks within selected CDs.

Probability sampling methods were also used to conduct surveys of businesses such as the monthly retail surveys. In Australia, the first business surveys using probability sampling were conducted in 1947 and in 1956 in NZ. Sampling methods were also being used in the Australian Census to conduct post-enumeration surveys and to support quality control of Census processing.

Copyright © 2023 Dennis Trewin. Published by [International Association of Survey Statisticians](#). This is an Open Access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2 Developments in Household Survey Methods

Household surveys enabled a massive increase in social statistics in both countries. Following the 1971 Australian Population Census, a significant effort was made into developing a multi-stage area sampling frame that could be used for multiple surveys as well as providing a rotating sample for the main survey – the quarterly labour force survey which had started in the early 1960s and became monthly in 1978. A stratified self-weighting design was used. Households remained in the sample for eight consecutive surveys before being rotated out. Independent population age x sex benchmarks were used to adjust for under-enumeration which was particularly high for young male adults. The new design provided for a 40% reduction in variance without any increase in sample size.

During this time a program of special household surveys was developed in both countries on topics such as household expenditure and income, health, time use, disability and education.

The enumeration methods have also evolved over time. Fifty years ago, face to face interviewing was exclusively used. Telephone interviewing was introduced in Australia in 1994 but a face to face interview was still used for the first time households were in the survey in order to obtain approval to conduct future interviews by telephone and obtain contact details. Before implementation, considerable effort was put into measuring any accuracy impacts from the change in collection mode. The existence of an identified impact delayed the introduction of telephone interviewing, but it turned out to be largely a first survey effect. Computer-Assisted Paper Interviewing (CAPI) was introduced in 2004 also enhancing the availability of para data to assist with survey design. Telephone interviewing was introduced in New Zealand a little later but face to face interviewing remained the dominant interviewing method. The accuracy of telephone lists was never high enough to be seriously considered as a sampling frame. The possibility of respondents completing the survey over the internet was introduced in 2014 in Australia following analysis showing measurement effects were relatively small.

An address frame of residential dwellings was developed for the 2016 Australian Population Census. It proved to be quite accurate and has been updated since then largely using external lists such as Australia Post and machine learning methods with satellite imagery as the data source. Since 2018, this frame has been used for the household survey program enabling more sophisticated means of controlling sample rotation and overlap between surveys. SNZ retains its 2 stage area sampling method for household surveys.

Methods for making estimates of aggregates and their sample errors have changed considerably over time as methods developed and more benchmark data became available. Model assisted GREG estimators are now mostly used.

3 Developments in Business Survey Methods

In Australia, probability surveys were first used for the business surveys (e.g. capital expenditure) in the 1940s and a little later in New Zealand. In the early years, the major cause of inaccuracies were missing units on the sampling frames. There was no great source for updating the frame. In Australia, a major effort in the late 60s and the early 70s was the development of an Integrated Business Register, providing for the hierarchy of enterprise groups, enterprises and establishments complete with industry codes, that could be used as a frame for all economic censuses and surveys. Tax data became available to support updates but still required considerable follow-up effort before new businesses could be added to the frame with confidence. Similar developments happened in New Zealand but it was not until the mid-1980s that the required tax data access was available. Prior to then, field checks by CPI staff were commonly used.

The existence of a Business Register enabled the application of a method, referred to as collocated sampling in Australia, to provide for the rotation of sampled businesses after they had been in the sample for a designated period of time (e.g., 3 years) (Brewer et al., 1972). Furthermore, it enabled

the control of overlap across surveys. Independently, Statistics Sweden had developed a very similar method. A simplified version was used in SNZ.

All the traditional economic censuses (e.g., manufacturing census) were converted to sample surveys in both countries across the following decade. More recently, most economic surveys tend to be economy-wide rather than industry specific.

Mail was the dominant data collection method with telephone interviewing mostly used to follow up non-response. The exception was the monthly retail survey where most of the data collection for small businesses was by telephone. Consequently, it was the first application of Computer-Assisted Telephone Interviewing (CATI) in the ABS, introduced after an extensive study of possible mode effects. About the same time, in both countries, the very largest enterprise groups were profiled so that their enterprise group-enterprise-establishment structure matched their financial accounting arrangements to the extent possible and their data collection was managed by so called large business units.

Despite the considerable efforts to integrate the data collections that provide source data for the national accounts, the statistical discrepancies within the accounts were still larger than desired. In the late 1980s, there was an extensive investigation into the reasons in Australia. It was found that a major reason was the inconsistent treatment of the missing units on the frame (mostly new businesses), businesses that were no longer operating, and non-response. Among other things, this led to a delay in detecting a turning point in the important survey of capital expenditure impacting government fiscal policy at the time. Standard procedures were developed by Methodology, including the estimation of new business provisions by industry, which were provided to all the surveys by an organisational unit especially set up for this purpose. Further data confrontation methods across collections, using the national accounting framework, were also put in place prior to the finalisation of the national accounts. This was a form of output editing but looking at multiple collections simultaneously. It resulted in considerably more accurate national accounts. An important outcome of the investigations was Ministerial support for greater access to tax data. Similar issues in SNZ led to proposals for a more systematic approach to economic collections.

4 Population Census

Statistical quality control methods for processing were first put in place for the 1961 Australian Population Census. Acceptance sampling methods were used which involved the acceptance or rejection of coded forms based on error counts determined by expert coders. Rejected lots were reprocessed. Studies showed that expert coders were not necessarily more accurate. At the 1976 Census, this was changed so the sample checks were used to provide information on the level and type of errors only. This information was used to identify ways in which the quality of processing could be improved e.g., retraining or improved coding instructions.

A Post-Enumeration Survey (PES) has also been conducted by the ABS since the 1966 Census. Its importance increased substantially following the 1976 Census. Following detailed demographic analysis, it was determined that Census counts were far more accurate estimates of the population if they were adjusted for undercount (at an age group x sex x State level) using PES data. Subsequently, the official population estimates for Australia have been adjusted using PES data after every Census. It was the first country to do so. The analysis also found, contrary to conventional wisdom at the time, the Censuses in the latter half of the twentieth Century were far more accurate than the first half. SNZ conducted their first PES in 1991. In SNZ, the PES is used for Census evaluation purposes only.

SNZ conducted surveys on Disability and Use of Maori Language using frames determined by responses to questions placed in the 2016 Census.

SNZ had field enumeration problems in the 2018 Census resulting in large non-response. They used administrative data to 'rescue' their 2018 Census supported by major methodological work. Based on this positive experience with the use of administrative data, the 2023 census is being designed

to make more use of this data and thereby reducing the reporting load on the public and possibly improving the accuracy of some aspects.

5 Increasing Interest in Non-sampling Errors

There has been long standing interest in both countries in measuring, understanding and controlling non-sampling errors. As part of the effort to control non-sampling errors, pilot testing was a standard procedure for new surveys or when introducing new methods. Until relatively recently, there was always a strong emphasis on maintaining high response rates to minimise non-response bias. They remain high by international standards although now survey designs focus more on ensuring samples are sufficiently representative of the population by using techniques such as adaptive sampling.

A research study into editing for the Retail Census showed that it introduced more errors than it discovered as the editing clerks learnt how to game the editing system so that each record passed the computer edits. Studies of other collections were consistent with this finding. The resources devoted to editing across all collections were considerable and did not contribute much to overall accuracy. These studies led to the introduction of more cost-effective macro-editing approaches that focussed on the most significant errors reducing costs as well as improving accuracy.

The increased effort into understanding and quantifying measurement errors led to consideration of the optimisation of Total Survey Error (TSE) rather than sampling error. This was inspired by Dalenius' work (1967) and preceded the more recent interest in TSE. In the ABS, it was applied to the design of the new Construction Industry Survey (see Linacre and Trewin, 1993) where one important decision from the TSE work was to use a more expensive field enumeration method for smaller businesses but with a smaller sample size. Subsequently, it has been recognised in design work that there are far better returns from methodological investments in frame maintenance and other non-sampling errors than clever work on sample designs.

This research also showed the importance of good management to reducing non-sampling errors. It was not just about design (see Trewin, 2001).

6 Use of Administrative and Big Data

In recent decades, there has been many innovative uses of administrative data. It has always been used to compile statistics in subject fields such as Foreign Trade and Births Deaths and Marriages. It has also been used to provide proxy indicators for compilations like the national accounts. Furthermore, it has been used to develop and maintain sample frames and benchmark data to help improve the efficiency of sample surveys. In more recent years, innovative uses include:

1. Data substitution (tax data). Considerable effort has been put into maintaining a good and trustworthy relationship with the Tax Offices. Access to tax data has increased considerably over time especially with the introduction of a Goods and Services Tax in both countries which provided monthly and quarterly data. One important use was data substitution. Studies showed the tax data was reliable (perhaps more reliable than data collected by the ABS and SNZ) especially if edited for the more significant anomalies such as coding errors.
2. Linked Data Sets. The links may be between two administrative data sets or between administrative data and Census/survey data sets. This has resulted in the creation of new richer data sets for the production of official statistics and supporting research;
3. Longitudinal Data Sets. A specific application has been the creation of longitudinal data bases; and
4. Big Data. There have been no real applications to date but its use in small area estimation is being actively investigated.

7 Longitudinal and Linked Data Sets

Neither the ABS nor SNZ have conducted many longitudinal surveys, but they have provided support to other agencies. However, in recent years administrative data has been used to create longitudinal data sets, sometimes using their own data sets, using linkages at the individual level across data sets. For example, Business Longitudinal Data Bases have been created in both countries. SNZ has created an Integrated Data Infrastructure combining administrative data for individuals.

As an example of longitudinal data sets involving survey data, longitudinal data files have been created from the monthly labour force survey taking advantage of the fact that 7/8th of the sample is common from one month to the next. A longitudinal data base of Census records has also been developed in Australia using statistical matching techniques. Starting with the 2006 Population Census, a 5% sample of Census records was retained without name and address identifiers but with sufficient information to allow statistical matching across the individual data sets. It also enabled linking across Censuses thereby establishing a longitudinal data set. Linkages with Death Records has enabled much more detailed morbidity analysis including for Indigenous persons in both countries.

Even when linking variables are available, they are subject to error or linking data not being specified consistently. Therefore, the development of algorithms to maximise the accuracy of linking has become a very important job for methodologists.

8 Researcher Access

Among the major changes to ABS and SNZ legislation in the 1980s were legal provisions to enable them to release unidentifiable microdata. This provided constraints on access which some researchers found too limiting. Recognising inadequate use of microdata has high costs, in the early 2000s data laboratories were introduced where researchers could work in a safe setting with supervision and checks to ensure the confidentiality requirements were met. This was later extended to use of Remote Access Data Laboratories so that it was not necessary for researchers to visit the Statistical Offices.

It is also important that researchers have good quantitative knowledge of measurement and other errors so they can be taken into account in the analysis (see Biemer and Trewin, 1997).

A more recent development has been ABS and SNZ becoming custodians for linked data bases data sets (including links with some of its own data sets). Data laboratories are often the only way to access these valuable data sets.

9 Non-ABS and Non-SNZ Surveys

Increasingly, surveys are being conducted by other government agencies. Following the introduction of the 1975 Statistics Act in New Zealand, a survey control function was introduced. Every proposed survey by other agencies had to be submitted to the Minister of Statistics for approval of the sample and survey design. SNZ did the analysis necessary to make a recommendation to the Minister. At the request of Government, the ABS introduced a similar function in 1997. The emphasis was very much on ensuring these surveys were fit for purpose rather than a design that was up to the standard of the official statistical agency.

10 The Future

The role of the survey methodologist has changed massively over the 50 year period. It is no longer sufficient to be an expert in sample design. Mixed mode data collection techniques will become very prevalent generating new methodological challenges. Sample surveys are only one source of data for official statistics and are often used in combination with other data sources. For example, further use of administrative data and big data (e.g. scanner data and satellite images) can be expected, sometimes involving machine learning applications. This would include linked data sets.

Maintaining the quality of surveys to be fit for purpose will be a big challenge. Non-response is already a big issue and will become even more of a challenge as will the maintenance of good quality frameworks. Sample designs and methods that adjust for these types of deficiencies will grow in importance. The demand for data from researchers will increase requiring the development of methods to improve access with confidentiality protection arrangements that meet public scrutiny.

11 Acknowledgements

I would like to acknowledge the very significant contributions made by Siu-Ming Tam (formerly ABS) and Vince Galvin (SNZ) and Len Cook and Paul Maxwell formerly of SNZ.

References

- ABS (2005), *Informing a Nation: The Evolution of the Australian Bureau of Statistics*, ABS, Canberra
- Biemer, P and Trewin, D. (1997), A review of measurement error effects on the analysis of survey data, in *Survey Measurement and Process Quality* by L. Lyberg et al, pp 603-632, John Wiley & sons, New York.
- Brewer, K., Early, L.J., and Joyce, S. (1972). Selecting several samples from the same population. *Aust. J. of Statist.*, 14(3), 231-239
- Dalenius, T. (1967). Nonsampling errors in census and sample surveys. *Report no 5 in the research project Errors in Surveys*. Stockholm University.
- Linacre S., and Trewin D. (1993), Total Survey Design - Application to a Collection of the Construction Industry, *Journal of Official Statistics*, 9(3), 611-621
- Trewin, D. (2001), The Importance of a Quality Culture, *Statistics Canada International Symposium Series: Proceedings 2001*.