

Methodological News

A Quarterly Information Bulletin



ABS Methodology and Data Management Division

September 2011

Articles

Getting a handle on individual diversity within areas	2
Exploring methodologies for estimating energy consumption	3
Google's Chief Economist visits the ABS	4
Industrialisation of Statistical Processes, Methods, and Technologies	5
Design and Estimation for the Agricultural Land and Water Ownership Survey	6
How to Contact Us and Subscriber Emailing List	7

Getting a handle on individual diversity within areas

Analytical Services Branch (ASB) has recently completed and published a study on individual diversity within areas, as part of its ongoing research on Socio-economic indexes for areas (SEIFA).

SEIFA seek to summarise the socio-economic conditions of an area using relevant information from the Census of Population and Housing. The SEIFA indexes are widely used measures of relative socio-economic advantage and disadvantage at the Census Collector District level. The indexes provide contextual information about the area in which a person lives, but within any area there are likely to be individuals with different characteristics to the overall population of that area. If inferences are made about these individuals based purely on the characteristics of the area in which they live, they could be misleading and there is potential for error in any conclusions – this is referred to as the ecological fallacy.

Using 2006 Australian Census of Population and Housing data, ASB explored individual level diversity within areas by creating and analysing two person-based socio-economic indexes: one of relative disadvantage and the other of relative advantage and disadvantage. The conceptual and methodological basis for these indexes was established by Baker and Adhikari (2007). The primary purpose of the study was to illustrate how individual level index scores can be used to illustrate and measure the diversity of socio-economic advantage and disadvantage within area level SEIFA. Secondary to this analysis of diversity, the

research served to highlight the advantages of SEIFA when compared with individual level indexes of relative socio-economic advantage and disadvantage, including maximising the proportion of the population receiving an index score.

Results showed that care that needs to be taken when using SEIFA information to draw conclusions about individuals who reside in those areas. Analysing the two individual level indexes of socio-economic advantage and disadvantage constructed in the study has facilitated the appraisal of diversity within areas, something that has until now not been possible except in Baker and Adhikari (2007). The results were important illustrations of how diversity of advantage and disadvantage within an area can exist, and the extent to which individuals with differing levels of socio-economic advantage and disadvantage reside in the same area.

The study has made clear some of the shortcomings of individual level indexes, and why SEIFA remains an important, robust product. Firstly, concerns with substantial population exclusions limit the applicability of the analysis; approximately one-third of the population counted in the 2006 Census were excluded from the individual level index construction process for applicability, compared to 0.6% of the population excluded for 2006 SEIFA. This vast difference reflects the robustness of the SEIFA indexes, namely that it maximises the proportion of the population that receives an index score. SEIFA is also more theoretically and conceptually sound because it is based on variables chosen for applicability in an area-based index, it is externally validated, and the aggregate nature of the data and

stringent exclusion rules both work to ensure that there is sufficient meaningful data in an area to support index construction.

Given the analysis presented in the research, users of SEIFA will understandably be wondering if a product can be released that enables them to appropriately tackle the issues involved. Before attempting this, the following critical issues would have to be resolved: consensus on the definition of individual level advantage and disadvantage, the best set of variables to measure it, a means for validating individual level indexes and a strategy for integrating an individual level product with the existing SEIFA indexes to ensure that it is a useful addition. ASB is recommending that the analysis in the research paper be repeated after the release of SEIFA 2011 because of the introduction of the Australian Statistical Geography Standard (ASGS). There is an expectation that the ASGS will better capture the socio-economic gradient within areas. In the meantime, ASB recommends using the SEIFA indexes for socio-economic analysis, bearing in mind the caveats relating to those measures not being attributable to individuals, but only to the average relative socio-economic advantage and disadvantage in an area.

The published paper can be found on the ABS website, catalogue number [1351.0.55.036](#).

References:

Baker, J. and Adhikari, P. (2007). "Socio-Economic Indexes for Individuals and Families", Methodology Advisory Committee Papers, cat. no. 1352.0.55.086, Australian Bureau of Statistics, Canberra.

Further Information

For more information, contact [Phillip Wise](#) on (02) 6252 7221 or [Rosalynn Mathews](#) (02) 6252 5257.

Exploring methodologies for estimating energy consumption

The Analytical Services Branch is currently working with the ABS Energy Statistics Unit on a project that aims to develop a methodology for estimating the energy consumption of businesses which are not required to report energy consumption under the National Greenhouse and Energy Reporting System (NGERS).

Australia's national energy statistics (Australian Energy Statistics or AES) was produced annually by the Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), but recently moved to the Bureau of Resources and Energy Economics (BREE). Before 2010, AES was based on data from the Fuel and Electricity Survey (FES) and other sources. From the 2009-10 reporting year, FES was replaced by the National Greenhouse and Energy Reporting System (NGER) as the main collection instrument.

The National Greenhouse and Energy Reporting Act 2007 (the NGER Act) requires all businesses that are a constitutional corporation and meet a reporting threshold for greenhouse gases or energy use or production for a reporting (financial) year to register with the Greenhouse and Energy Data Officer, Department of Climate Change and Energy Efficiency. The reporting threshold for 2009-10 was 350 terajoules (TJ) dropping to 200TJ in 2010-11.

This means the energy consumption of many Australian businesses, across different industry divisions, will not be captured by NGERS data as they will fall under this reporting threshold. This undercoverage is required to be addressed to get a better picture of overall energy consumption.

The ABS has undertaken some initial analyses of the possible level of undercoverage using results compiled from the 2008-09 Energy, Water and Environment Survey (EWES).

The ongoing research in ASB points to several options for estimating the energy consumption gap. The modelling approaches have their theoretical foundation from production economics and applications from empirical estimations of production and cost functions. Data at both unit record and industry level can be considered which will have impacts on the choice of modelling methods.

Further Information

For further information, please contact Kay Cao on (02) 6252 5264 or kay.cao@abs.gov.au

Google's Chief Economist visits the ABS

On ABS invitation, Professor Hal Varian, the Chief Economist of Google, visited the ABS in early July. Professor Varian's visit coincided with a trip to Australia for the Australian Conference of Economists in Canberra, where he was a keynote speaker.

During his visit, Professor Varian gave a seminar, titled "Predicting the Present with Google Trends" at ABS, and met with Brian Pink and senior executives. In his seminar, Professor Varian illustrated how query data may be used to predict economic activities. He explained how Google Trends provide daily and weekly reports on the volume of queries related to various industries and this query data may be correlated with the current level of economic activity in given industries and thus may be helpful in predicting important economic indicators. He discussed the research tools used in Google, such as Kalman techniques, Bayesian Model Averaging, Spike and slab and horizontal regression. Professor Varian also contrasted privately owned high-frequency/real time data with government agencies' accurately constructed but low frequency data, and discussed the challenges around integrating private sector real-time information with traditional government statistics.

The meetings between Professor Varian and ABS involved discussions around Google's approach and experience with data management, and created potential for future collaboration with Google on data management issues.

The visit was a good opportunity for the ABS to discuss the experiences of Google and it commenced what can hopefully become an ongoing dialogue between the two organisations.

Further Information

For more information about Professor Hal Varian's visit to the ABS or the seminar presented at the ABS, please contact Gokay Saher on (02) 6252 6209 or gokay.saher@abs.gov.au

Industrialisation of Statistical Processes, Methods, and Technologies

At last month's International Statistical Institute 2011 World conference, the ABS presented a paper on the topic of the industrialisation of statistical processes, methods and technologies. This paper was a collaboration between IMTP (Geoff Lee) and MDMD (Lisa Apted, Philip Carruthers, Daniel Oehm and Frank Yu) and discussed the industrialisation of statistical activity in national statistical agencies such as the ABS. The paper illustrated three examples within the ABS where industrialisation might be usefully applied. This article briefly summarises the paper.

During the industrial revolution, so-called "cottage industry" approaches were replaced by standardised processes, new methods, and innovative technologies which had far reaching impacts on the efficiency and effectiveness of manufacturing production. The revolution created an environment in which outputs could be delivered more easily

and cheaply. The outcome was not only a reduction in the resource base required but also an outpouring of new possibilities for utilising the products made possible by their increased availability at reduced cost. The "industrialisation" concept of standardised production processes -- where new production methods and technological advances allowed production on a large scale -- worked in a manufacturing context, and there are sufficient similarities to suggest it should work equally well in a statistical production context. Here, it is assumed that cheaper access to greater quantities of high quality statistics will encourage the generation of more innovative presentation, reporting and analysis methodologies.

In the official statistical context, it is important to understand that industrialisation does not represent the robotic automation of standardised statistical activity. Rather, it offers improved capacity to replicate basic processes, freeing up analyst resources so they can add value where needed, with the support of knowledge-based decision-making tools.

There are many different statistical processes within the ABS that are amenable to greater industrialisation, including the seasonal analysis of time series, microdata confidentialisation and the compilation of price indexes. Standardised metadata can be used to drive business processes and capture data relationships to promote greater harmony between similar or dependent activities. Since at their core all of the production processes for official statistics are about manipulating and quality assuring information, information management must

be a strategic focus to achieve greater harmonisation and "industrialisation".

Anyone who has been involved in the process of developing an international statistical framework will be aware that it can be a slow and tedious process. Reducing the rate of innovation in the production methods of official statistics is a genuine risk to the "industrialisation" philosophy. It need not necessarily be so, as the experience of the growth of the internet has shown. The internet has created an environment within which enormous bursts of technological creativity and innovation occur - and yet at its heart it is reliant upon some basic standards which have been widely adopted. The task for methodologists working in the production of (official) statistics therefore becomes one of establishing which methods can and should be standardised, and which should not.

Further Information

For more information, please contact Philip Carruthers on (02) 6252 5307 or philip.carruthers@abs.gov.au

Design and Estimation for the Agricultural Land and Water Ownership Survey

Readers may be aware that foreign ownership of agricultural land in Australia has been the subject of considerable media and community interest in recent months. When this issue was raised in Parliament late last year, Treasury officers approached the ABS with a request to conduct a survey to inform the Government on the degree and extent of

foreign ownership of agricultural businesses, land and water. This request was passed on to the Environment and Agriculture Business Statistics Centre (BSC) who were asked to commit to a release date of September 2011. Anyone with experience of survey development and conduct will appreciate that this is a very tight timetable, especially in a year when the BSC is conducting the five-yearly Agricultural Census.

The outputs desired are business counts, land area and volume of water entitlement, by degree of foreign ownership of the business (5 categories) and at national and state level. These are to be produced from a sample of about 10,000 from a frame of 162,000 units. The survey presented the BSC with a number of challenges from the outset:

- the degree of foreign ownership can be difficult to determine where there are a number of businesses in the "ownership chain", or in cases of joint ownership
- there are regional differences in the way water use is regulated
- the frame used (the preliminary frame for the Agricultural Census) comprises all Australian-registered businesses with known agricultural activity. This is the only practical frame that could be used, but requires us to overcome the challenge posed by farms that lease land from non-agricultural businesses which are not on the frame.

Sample design also had a number of challenges:

- there is little relevant data to guide sample allocation (a survey of foreign ownership in the agricultural sector was last conducted in 1984)
- the businesses of interest are a sparse sub-population (only about one in 160 businesses is expected to have some degree of foreign ownership)
- the only variable available for sizing units does not correlate very well with land area

Initial work by the BSC identified a small number of units thought to be foreign-owned and the distribution of these units over industry and geographical region was used to inform the stratification. There was no size component of the stratification other than a separation of units into the completely-enumerated and sampled sectors. Instead, the importance of larger units was captured by making the selection probability proportional to a measure of the unit's agricultural production. Unconditional Poisson selection was used to select the sample. Since this means that sample sizes are not fixed, weights were calibrated to post-stratum populations in order to improve the accuracy of estimates and variances.

The publication was released on September 9, the culmination of a significant effort by the BSC and its MDMD support to meet a non-standard request in a difficult time frame.

Further Information

For more information, please contact Jos Beunen on (02) 9268 4647 or jos.beunen@abs.gov.au

How to Contact Us and Subscriber Emailing List

Methodological News features articles and developments in relation to methodology work done within the ABS Methodology and Data Management Division. By its nature, the work of the Division brings it into contact with virtually every other area of the ABS. Because of this, the newsletter is a way of letting all areas of the ABS know of some of the issues we are working on and help information flow. We hope the Methodological Newsletter is useful and we welcome comments.

If you would like to be added to or removed from our electronic mailing list, please contact:

Valentin M. Valdez
Methodology & Data Management Division
Australian Bureau of Statistics
Locked Bag No. 10
BELCONNEN ACT 2617

Tel: (02) 6252 7037
Email: methodology@abs.gov.au