Australian Bureau of Statistics

METHODOLOGICAL NEWS

A QUARTERLY INFORMATION BULLETIN FROM THE METHODOLOGY AND DATA MANAGEMENT DIVISION

September 2007

National Statistical Service (NSS) Leadership Branch

The National Statistical Service Leadership Branch is made up of two Sections: Statistical Coordination (Statistical Clearing House (SCH) and National Statistical Service (NSS) teams) and the National Data Network (NDN) Business Office.

The Statistical Clearing House is responsible for reviewing all surveys involving 50 or more businesses conducted by or on behalf of Australian Government departments and agencies. The primary purpose of the SCH is to reduce the burden of Australian Government surveys on businesses by ensuring such surveys do not duplicate existing collections and are of sufficient quality to warrant the burden imposed. The Clearing House is celebrating its 10th birthday in October 2007.

The National Statistical Service (NSS) team has a particularly important role in driving overall NSS strategy taken by the ABS and providing leadership on key cross cutting projects. Specifically, the area:

- assists the development of NSS policy and strategy;
- markets and promotes awareness of the NSS;
- leads cross-cutting NSS projects (including pilot projects for the National Data Network);
- promotes the development of tools and skills both within the ABS and across the NSS; and
- supports associated initiatives and forums (such as the Community of Users and Producers of Statistics (CUPS) and the Australian Government Statistical Forum).

The NDN Business Office provides secretariat support for the NDN Interim Governing Board and develops policy proposals for Board approval. In addition to their secretariat role, the Business office also undertakes:

- provision of business analysis for potential NDN clients;
- managing the NDN website, which is the main communication channel to users, custodians and interested organisations;
- production of the monthly NDN newsletter;
- managing the ABS Node including all registrations for the NDN;
- provision of specifications to the development team for NDN technical development; and
- production of documentation including business rules, participation agreements, metadata requirements, manuals and help support.

If you would like to know more about NSS Branch activities, please contact Mark Lound, Director, Statistical Coordination on (02) 6252 6325 or John Sant, A/g Director, NDN Business Office on (02) 6252 5717.

The 2011 Census - Your chance to have a say

While continuing to deliver more new and exciting products and services from the 2006 Census, the ABS is turning its attention to the next Census of Population and Housing to be held in August 2011.

The ABS will soon be inviting public comment on the content and procedures of the next Census. The invitation to have a say in the way the nation's largest statistical collection is undertaken is in the Information Paper: 2011 Census of Population and Housing: ABS Views on Content and Procedures to be released on 15 October 2007.

This is the first in a series of information papers about the 2011 Census. It marks the first step in the public consultation process for 2011 and outlines ABS proposals for the next Census, including:

- procedures for conducting the Census
- arrangements to protect the privacy of individuals
- measures to ensure the confidentiality of the information collected
- topics to be included
- topics under review, including new topic proposals
- topics to be excluded

How do I make a submission?

Submissions can be lodged either online, electronically by email or in hardcopy. The submission period will open on 15 October 2007 when the Information Paper is released and will close on 31 March 2008. Users of Census data and interested members of the public are

invited to make submissions on any aspect of the Census.

The Information Paper and Submission Form will be available on the ABS website (for free download) at www.abs.gov.au/2011censusviews from 15 October 2007. Information and guidelines about making a submission will also be available from the website.

Information Sessions on 2006 Census products and 2011 Census Topic Consultation are planned for each capital city during October and November 2007. These meetings are open to all interested members of the public and provide an opportunity to hear more about plans for the next Census, meet ABS staff and ask questions.

In addition to providing copies of the Information Paper to people and organisations with an interest in Census data, the ABS will notify the public of the release of the Information Paper through:

- advertisements placed in major newspapers across the country
- media releases
- the ABS website

What happens next?

Following the assessment of submissions and consultation meetings, recommendations on the content and procedures of the 2011 Census will be discussed with the Australian Statistics Advisory Council later in 2008. The ABS will then prepare a submission to the Federal Government during 2009. A final decision on the 2011 Census topics is expected to be made by the Federal Government in late 2009. The ABS will then release a further Information Paper outlining the final nature and content of the 2011 Census.

All individuals or organisations who have made submissions will be advised of the final outcomes.

For any further information, please contact Anabel Thomasse on (02) 6252 5751.

ABS Using Operations Research to Improve Efficiency and Effectiveness of Business Surveys

The ABS and the world around us is becoming more complex. Huge numbers of choices and relentless time and cost pressures make the decisions we face more daunting and more difficult. Meanwhile, our systems are generating massive amounts of data about how we do our business providing us with opportunities for independent and objective evaluation of our processes. For example, the introduction of a centralised Provider Management Information System (PIMS) for recording all contacts made to providers in Business surveys has given us the opportunity to objectively analyse and improve the effectiveness of intensive follow-up (IFU) processes. The Operations Research Unit (ORU) was established within the Statistical Support Branch in July last year and is tasked with taking advantage of this and other similar opportunities to answer questions such as:

- Which strategies result in the highest rate of response for the lowest amount of cost?
- What should we do differently and how should we do it?
- Where are we expending a lot of effort without receiving a lot in return?

Investigations to date have centred around the annual Economic Activity Survey (EAS) and two sub-annual surveys - the Average Weekly Earnings Survey (AWE) and the Quarterly Economy Wide Survey (QEWS). Generally the IFU strategies for these and other collections involve the use of reminder letters and follow-up phone calls. However, the number and timing of these letters and phone calls relative to the due date varies considerably. For example, some collections use pre-approach letters, some quarterly collections have up to four reminder letters, while some annual collections may only have two. The length of the IFU cycle varies from collection to collection and the overall response rate or imputation rate required by each collection also differ. Within each collection, contact of providers is usually prioritised by significance. This significance is assigned based on different factors for each collection. Providers are called randomly throughout the day without regard as to whether some times are better to call particular providers or not.

Investigations into EAS, QEWS and AWE contact data have identified a number of patterns that can be used by the Provider Contact Unit (PCU) to improve effectiveness of the IFU processes. For example, it was found that:

- towards the end of the cycle, a large amount of effort is expended sometimes without obvious gain, some providers are contacted up to 30 times without securing a response. In particular, in the AWE survey, calling providers more than 5 times takes 10% of the overall effort in terms of calls and results in an increase to the non-imputed contribution of estimates of 3% and an increase in form receival rate of 3%. Being able to identify a point in the IFU process after which further gains are expected to be very small will enable to develop more effective strategies for following up outstanding providers that do not respond to simple intensification of effort;
- best time of day to contact a provider varies between industries. This information can be used to schedule provider follow-up in a way that maximises the chances of making a contact with the provider. In particular it appears that a higher rate of contact for businesses in Education could be achieved by avoiding the 1-2pm time slot;
- there is a lag between a follow-up attempt and a response triggered by that attempt. In the annual survey of EAS, the lag was around 14 days whereas for the quarterly survey of QEWS, the lag was around 6 days. By refraining from further contact with the providers during this lag, the provider load and follow-up effort can be reduced without negative impact on the response or imputation rate;
- providers are less likely to respond if a long period elapses between follow-up attempts. This is not usually a problem in QEWS, but not contacting

someone for over 30 days can be a problem in EAS. By concentrating the follow-up effort in a way that ensures that all providers are followed-up within a reasonable time from the last attempt better response and imputation rates can be achieved earlier in the IFU period and number of providers that need a large number of contacts at the end of the IFU period can be reduced;

- around 93% of providers who responded without phone follow-up in one quarter respond without phone follow-up in the subsequent quarter. The PCU has used this information to minimise unnecessary contact with "good" providers.

Analysis of existing contact data is not sufficient to determine whether proposed modifications to follow-up strategies actually result in expected improvements and to ensure that these changes do not compromise the quality of statistics collected. In order to do that, carefully controlled experiments that allow alternative strategies to be tested and if necessary modified are required. Currently, two trials are planned to test alternative dates for phone follow-up in QEWS and EAS collections. These trials will allow the level of improvement to be quantified and if successful will enable evidence-based changes to IFU strategies for these collections.

Future work will concentrate on following up the results of the trials and expanding the research to more surveys. Further work is also planned on details relating to reminder letter timing, the use of pre-approach letters, extensions, significance flags and more control of call patterns.

For further information, please contact Louise Gates on (02) 6252 6540.

Acceptance Sampling Based Clerical Review in Probabilistic Matching

During the Census processing period, several quality studies are being conducted as part of the Census Data Enhancement (CDE) project. There are two types of quality studies, the first to assess feasibility and quality of linking without names and addresses and the second to help improve ABS statistical outputs. As part of these studies, names and addresses as well as other variables are being used to link Census data with other selected data sets.

The method used is probabilistic linking where the aim is to link records that are believed to belong to the same person from two different data sets. This method has its foundations in the method proposed by Fellegi and Sunter (1969). In this method candidate record pairs are given a weight based on the degree of agreement between fields on the two records. Record pairs with a weight above some upper cut-off are declared links while those with a weight below some lower cut-off are declared non-links. However, there are many record pairs that cannot be automatically assigned a status and are designated for clerical review. Clerical review involves human assessment of each record pair to resolve match status.

Clerical review is a time intensive stage of the data linking process requiring a high level of VDU and keyboard equipment use. Some linkages can generate thousands, ten or potentially even hundreds of thousands of record pairs for clerical review. The CDE project has implemented an acceptance sampling based approach to dramatically reduce the amount of clerical inspection.

Acceptance sampling is a well established statistical method that replaces 100% inspection with inspection of samples selected from batches. The clerical review pairs are ordered by linkage weight and divided into batches . A sample is selected from each batch and each selected record pair is inspected. The number of matched and non-matched pairs in this selection is compared against a set of critical values. The entire batch is sentenced on the basis of these comparisons. If the number of matches observed is less than the lower critical value then the batch is assigned as non links. If the number of matches observed is greater than the higher critical value then the batch is assigned as non links. Otherwise the batch is assigned for complete manual clerical review.

As sampling is used there are a number of risks. Batches containing a high proportion of actual matches that would be assigned as links may, by chance, be sent for clerical review. Similarly, batches containing a low proportion of actual matches that would be assigned as non-links may also be sent for clerical review. Both these cases would result in wasted effort. Lastly, batches that should be clerically reviewed because they contain a relatively high proportion of both of true matches and true non-matches may, by chance, be assigned as either links or non-links and not reviewed in full. However, careful selection of a clerical review threshold and sample size enables these risks to be quantified and controlled. Acceptance based clerical review has provided an accurate and reliable means of assessing and setting the most appropriate clerical review bounds.

The acceptance sampling software, incorporated into the linkage software, FEBRL (http://datamining.anu.edu.au/software/febrl/febrldoc/) by ABS's Statistical IT Facilities, is flexible and user-friendly. It allows the operator to set sampling parameters, move freely through the batches, override the automatically assigned batch status and manually assign the status of a batch or group of batches.

Using this method we have been able to reduce the number of record pairs for clerical review from 11,000 to 4,000 in one linkage.

For further information, please contact Tenniel Guiver on (02) 6252 7310.

Exploring the Link between Innovation and Business Performance

It is generally believed that innovation is a key driver of firm-level performance. However, for Australia, quantifying the relationship between innovation and performance growth (including productivity) had been difficult because there was no single data source that contained both innovation and performance measures. The complexity of the innovation process also complicated this type of analysis.

The Australian Bureau of Statistics (ABS) developed a new business innovation survey in 2003 as a result of strong user demand for updated and internationally comparable statistics covering technological and non-technological innovation. Using this data source, a recent study undertaken by the Analytical Services Branch (ASB) explored the effect of innovation on performance growth at the individual business level. This study used business-level data from the 2003 Innovation Survey, which was augmented with performance information from the ABS Economic Activity Survey (EAS) and the Business Income Tax (BIT) and Business Activity Statement (BAS) data of the Australian Tax Office (ATO).

This research applied the Crepon, Duguet and Mairesse (CDM) framework to explore the association between innovation and business performance. The CDM is a three stage model which includes the following relationships: the innovation investment equation linking the level (if any) of innovation expenditure to its determinants; the innovation output equation relating innovation input and other determinants such as firm, industry and market characteristics to innovation output; and the business performance equation which examines the effects of innovative output and other factors on growth in business performance.

This study was exploratory in nature and is part of the ongoing overall research into the determinants of innovation expenditure and innovation output, and of the impact of these on associated firm performance in Australia. The aim of this analysis was to inform future data collection and research direction, rather than to generate definitive results.

For more information, please contact Ruel Abello on (02) 6252 6307.

Making Quality Visible Update

In the March 2007 edition of Methodological News, we reported on the Making Quality Visible (MQV) initiative that was underway. This article is an update of the progress that has been made in recent months on various aspects of the MQV initiative.

Quality declarations

Quality Declarations (QDs) are statements on the quality of statistical releases that have been written specifically for web-based dissemination. They describe the quality of a statistical release using the six dimensions of the data quality framework adopted by the Australian Bureau of Statistics (ABS) - relevance, timeliness, accuracy, coherence, interpretability, and accessibility.

In the past few months, the ABS has committed to the creation and release of QDs accompanying statistical products on the web, and planning for implementation is well underway. The first QD will be associated with the second release data of the Census of Population and Housing, and will be released in October 2007. This QD will cover some of the important data quality aspects of the 2006 Census. The release of QDs will occur progressively from January 2008 onwards, with a majority of ABS statistical releases acquiring a QD in 2008.

General information on QDs is available on the ABS website, and supporting information for author areas in the form of a QD User Guide and a 'Reporting Quality in the ABS' course will become available in October.

Electronic Publication and Metadata Vision

The Vision aims to improve the way the ABS communicates data and metadata electronically to facilitate user discovery of information, assessment of fitness for purpose and understanding of the key stories within the data.

Progress has been made in the development of systems to cater for the electronic publication and metadata vision. Various statistical releases have moved towards an electronic magazine (eMagazine) design for the web. This design sees the use of 'chunking' of sections of information with the use of layering to present more detailed information progressively from a link in a previous layer. Releases that have gone the way of the eMagazine design include Australia's System of National Accounts, 5204.0, Qld Stats, 1318.3.

Quality Assistant

A Quality Assistant was released to ABS staff in September 2007. The Quality Assistant, available to staff from the ABS Welcome Page, is a one stop shop for anyone in the ABS trying to find information on quality initiatives in the ABS. The Quality Assistant will provide links to manuals, information on the Quality Infrastructure System, Quality Gates and Quality Measures along with other useful information.

Quality Infrastructure System

To provide ready access to a range of relevant and timely quality measures, the ABS has been developing a Quality Infrastructure System (QIS), an integrated infrastructure that will allow the capture, storage and use of quality measures over the end to end business processes across all ABS data collections. QIS is designed to make quality measures readily accessible for use in quality assurance processes, as well as for reporting on quality and making quality declarations about data.

The end of August saw a major milestone when QIS was released into production as a corporate tool. This first version release of the full production system includes output reporting tools developed using SAS

Business Intelligence (BI), and utilises 13 quality measures for economic surveys, which span form response rates, frame quality, measures of accuracy and adjustments to data. The current production version of QIS only has information from ABS Survey Facilities (estimation, imputation and outliering information) for one business survey, the Retail Business Survey. We currently have two other surveys trailing their data in the test version of QIS, with the view to move to production after sufficient testing. The vision is to expand the number of quality measures supported by QIS. The future roll out of QIS to more collections will be tightly linked with the development of quality gates to manage statistical risk, and will be determined in consultation with individual survey areas.

Implementation for population surveys will coordinate with SSG's Business Activity Monitoring system, with the production roll-out due to occur in October 2007.

For more information on this work, please contact Bruce Fraser on (02) 6252 7306 or Narrisa Gilbert on (02) 6252 5283.

Excel Form Standards Update - Multi-unit (Horizontal) Forms

The Australian Bureau of Statistics (ABS) has been using the Microsoft product Excel to design self-administered business forms for many years and standards for the design of these forms were released in 2002. The standard forms were designed to look very similar to their standard paper form equivalents with a few exceptions relating to the electronic mode, for example using Arial font rather than Times.

Recently an audit of electronic forms used for business survey respondents was conducted to understand the current situation in order to manage the growth and nature of e-reporting to the ABS in the future. A key finding from the audit was that there were some quite extreme departures from ABS Excel form standards to meet the requirements of the surveys and reduce the burden of the respondent.

Form design standards have therefore been developed for a new type of form, which was named a 'Multi-unit (horizontal) Excel' form. The existing form type was named a 'Conventional (vertical) Excel' form. Multi-unit Excel forms have a spreadsheet-like layout. These forms have particular application to surveys that require multiple responses, that is, for different business units, on the same form.

Each row in the spreadsheet potentially contains the response for an individual business unit. The forms are also sometimes used where the respondent wants to report on a by unit basis, and the ABS is interested in the aggregate. Typically, each survey question is in a column heading and each row represents a responding unit or a repeated data item broken down in some way e.g. information from different states in Australia, or different departments in a university.

Generally the Excel standards, for example the use of colour, lines, font, headings and labelling still apply for Multi-unit forms. These forms are also still required to have all the standard components, i.e. Front, General instructions and How to submit pages.

The major difference between the form types is the layout, with Conventional forms having a vertical layout, i.e. questions being listed down the page in the same arrangement as the paper form equivalent, and the Multi-unit having a horizontal layout with questions listed in each column across the spreadsheet. Other variations include allowing smaller font sizes for the Mutli-unit forms to allow more space for the question text, and in some circumstances paraphrasing of questions due to space restrictions. Where any notes or questions are summarised, a separate pdf of the form and instructions should also be provided to the respondent.

Multi-unit Excel forms will allow the ABS to be more flexible when designing electronic forms, and can greatly reduce the time and effort required for respondents to provide their data.

For more information, please contact Leone Wardman on (02) 6252 7883.

The Children and Youth Statistical Portal

The Children and Youth Statistical Portal (CYSP) is the current pilot project for the National Data Network (NDN). The CYSP will provide access to a range of children and youth information resources from a single location. Catalogue entries containing discovery metadata, including a web link to these information resources will be available via the Portal after the software release scheduled for 8 October 07.

Developments underway

a. Approaching custodians

The ABS is currently approaching custodians of children and youth information resources to confirm their contributions to the Portal. As part of the Pilot Phase, a selection of contributing national and state agencies and organisations will be provided with Lite node software to enable them to load catalogue entries. To support custodians with this work, an Information Pack has been developed and distributed. The Pack is available from <http://www.nationaldatanetwork.org/NDN/NDNHome.nsf/home/Bookshelf?OpenDocument>.

b. Portal Discussion Forum

The Portal software scheduled for release in October 07 will include a Discussion Forum. This Forum will be moderated and users will need to be registered to 'post' to the Forum, with moderated 'posts' publicly available. Opening issues may relate to:

- integration of agencies within the children and youth arena,
- the usability of the Portal, and
- current work in the area of childrens' learning and development.

CYSP Reference Group meeting

A Children and Youth Statistical Portal Reference Group meeting was held on 1 August 07. This group was well attended, with representatives from a range of Australian and State/Territory Government departments and national organisations. Representatives were encouraged to engage with the Portal and contribute to the development through provision of catalogue entries of publicly available information resources.

For those interested in more information about the Portal contact Lorraine Cornehls on (02) 6252 6079). If you would like more information about the NDN in general, please call Patricia O'Reilly from the NDN Business Office on (02) 6252 5875.

Australian Economic Indicators Increases Web Presence

The Australia Economic Indicators (AEI) area has expanded its product range with the release of the Key National Indicators dashboards in August 2007. Dashboards are a series of short graphical representations that are designed to work like a vehicle dashboard ie. give the user a quick summary of how things are going. The AEI team worked in conjunction with Web Publishing in order to develop this product as a "proof of concept" and will take responsibility for updating the product on a daily basis as new data is released.

With web content being a major part of the Australian Bureau of Statistics output and reporting, the AEI area has shown leadership with this product, as a properly composed graphical summary is able to impart information more quickly than a series of tables. The development of the product took several months of work involving extensive contact and consultation with subject matter areas until consensus was reached as to the style, format and time series to be shown in the dashboards.

Web users can also now access data cubes of full length time series for all tables in AEI (cat. no. 1350.0). The full set of data cubes has been available since the May 2007 issue and this means that AEI will no longer publish the traditional "long time series" edition each February.

Feedback on the Key National Indicators dashboard is invited and any comments or suggestions should be provided to Jo Jackson on (02) 6252 6114.

How to Contact Us and Subscriber Emailing List

The Methodological Newsletter features articles and developments in relation to work done within the ABS Methodology and Data Management Division. By it's 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 welcome comments.

If you would like to be placed on our electronic mailing list, please contact:

Jayne McQualter
Methodology & Data Management Division
Australian Bureau of Statistics
Locked Bag No. 10
BELCONNEN ACT 2617

Tel: (02) 6252 7320
Email: methodology@abs.gov.au

Click on the following links to view the ABS Privacy Statement and Disclaimer
Privacy Statement | Disclaimer

Methodology & Data Management Division
Management Structure

Current at September 2007

First Assistant Statistician
Geoff Lee Tel: 5239

Assistant Statistician
Analytical Services Branch
Jill Charker Tel: 7290

Directors
Analytical Services

Glenys Bishop Tel: 5140

Jonathon Khoo Tel: 5506

Ruel Abello A/g Tel: 6053

Analytical Services

Joanne Baker A/g	Tel: 6992
Nicholas Biddle	Tel: 5061
Tetteh Dugbaza	Tel: 7221
Daniel Elazar	Tel: 6962
Tenniel Guiver	Tel: 7310
Anil Kumar	Tel: 5344
Peter Rossiter	Tel: 6024
Richard Solon	Tel: 5917
Getachew Tessema A/g	Tel: 5107

Assistant Statistician
Data Management & Classifications Branch
David Lengyel A/g Tel: 6430

Director
Data Management

Alistair Hamilton Tel: 5416

Census & SuperSTAR
Robert Brown Tel: 6211
Economic Statistics Support & Training
Philip Carruthers Tel: 5307
Metadata
Simon Wall Tel: 6300
Operational & Population Statistical Client Support
Margo Lockwood Tel: 5575

Director
Economic Standards & Classifications

Michael Meagher Tel: 7967

Economic Standards & Classifications

Barry Keeley	Tel: 7469
Celia Quiatchon	Tel: 5604
Jenny Foster	Tel: 6634
Linda Walker	Tel: 6252
Louise Perkins	Tel: 6790
Michael Robertson	Tel: 5826

Director
Population Statistics Standards

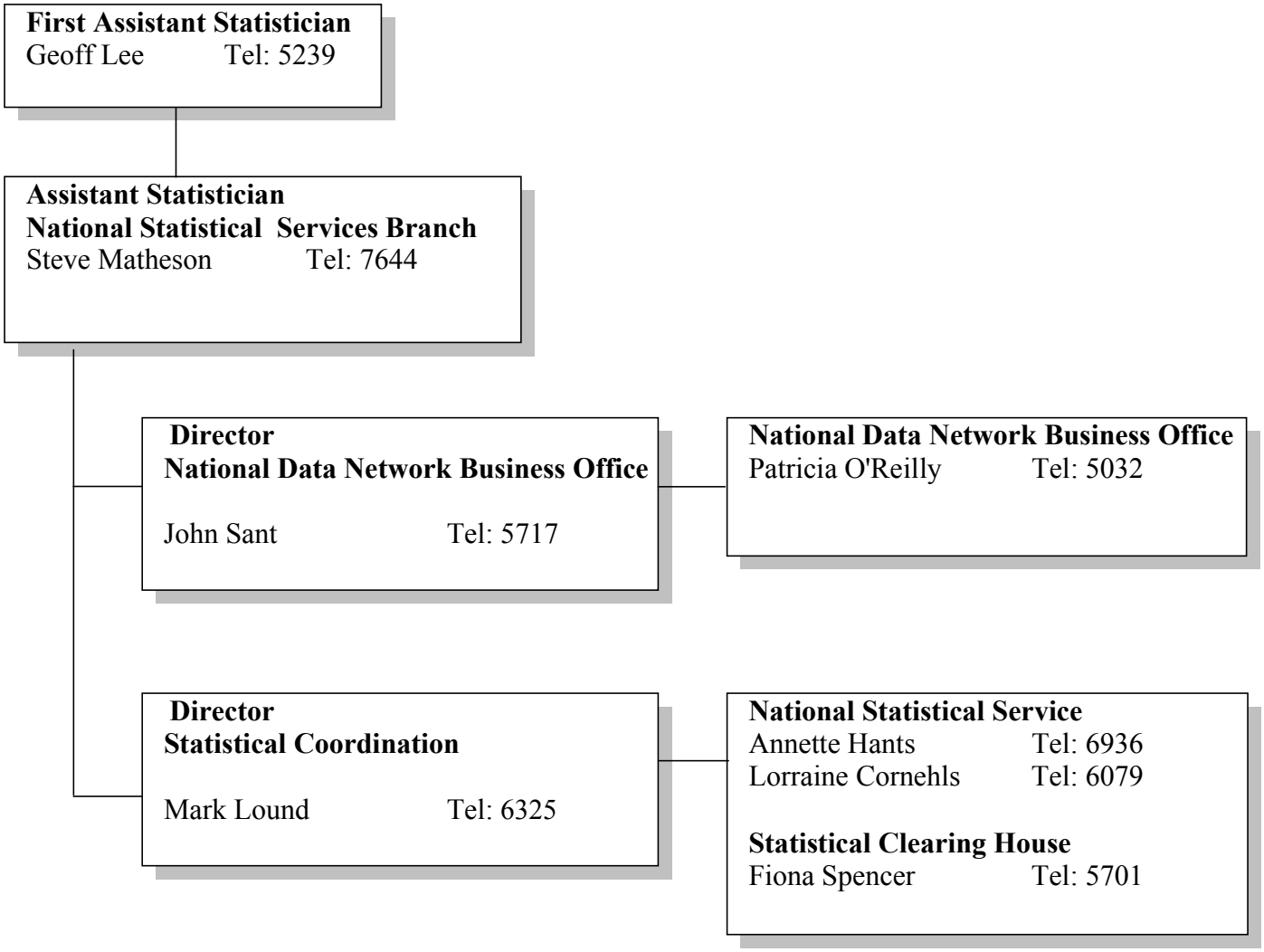
Joan Burnside Tel: 7074

Population Statistics Standards

Andrew Woolley	Tel: 7073
Rosa Gibbs	Tel: 7805
Tim Holloway	Tel: 6000
Tony Kershaw	Tel: 5453

Methodology & Data Management Division
Management Structure

Current at September 2007



Methodology & Data Management Division
Management Structure

Current at September 2007