



**Economic and Social
Council**

Distr.
GENERAL

CES/2005/13
22 March 2005

Original: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-third plenary session
(Geneva, 13-15 June 2005)

**USING BOTH INTERNET AND FIELD COLLECTION METHODS FOR THE 2006
CENSUS OF POPULATION AND DWELLINGS – THE NEW ZEALAND
EXPERIENCE SO FAR**

Invited paper submitted by Statistics New Zealand¹

BACKGROUND

1. New Zealand has a population of 4 million, located across a country with a land mass roughly the size of the United Kingdom. Around three quarters of the population live in the North Island, with the longer and larger South Island accommodating the remaining 1 million. The indigenous population (the Maori) makes up approximately 15 percent of the population, with 7.25 percent of the population being people of Pacific Island descent and 9.82 percent being Asian. The impacts of immigration are not uniform across New Zealand, with one in three Aucklanders born overseas, compared with the national figure of under one in five. Like many other similar countries, the population is ageing. The median age was 34.8 years at the 2001 Census of Populations and Dwellings, compared with 31.3 in 1991.
2. The Statistics Act provides that Population Censuses be conducted every five years in New Zealand. As with a number (albeit declining) of other countries, censuses in New Zealand are based on the model of enumerator drop off and collect, although a small proportion is returned by mail. Since 1945, separate dwelling and individual forms have been provided for all households. These forms have been available in both English and in a bilingual (English/Maori) format since the 1996 Census. Support for the Census remains

¹ Prepared by Brian Pink.

high, with under-coverage as measured in the Post Enumeration Survey, being only 2.8 ± 0.3 percent in 2001.

3. It is recognized that this high level of support is not something that can be taken for granted. New Zealand society is changing in many ways. Society is becoming more diverse, and people are busier, less available and less willing to accept the need to participate in statistical activities such as the Census. Public attitudes in New Zealand and in many other countries indicate a strong demand for choice and convenience when interacting with Government. This has been an important driver in deciding to proceed with the provision of an online form for the 2006 Census.

4. From a statistical perspective, the key drivers for an online form are:

- to improve/maintain participation in the census; and,
- improve aspects of census data quality.

Field tests that have been undertaken have proved both the feasibility of this approach and confirmed that the driver's objectives are being met.

5. As the 2006 Census will be the first mixed-mode census in New Zealand, the learning gained from it will position Statistics NZ to take advantage of the efficiencies that the technology can provide in the 2011 and following censuses. Statistics NZ has already made significant gains with imaging of forms since the 1996 Census, and would also be making savings in both collection and processing costs in the future. Other major benefits should include timelier and higher quality census data. There are also wider government benefits - the provision of online forms is an important contributor towards meeting eGovernment objectives for delivering a wide range of government services via the Internet.

6. The risks of inaccurately forecasting Internet take-up levels would place at jeopardy the integrity of the 2006 Census. There has therefore been no attempt to leverage efficiency gains in any of the traditional census processes. Paper forms are being printed for all households and individuals and there will be a full quota of field enumerators engaged for the 2006 Census. This is a longer-term investment in improving participation and the data quality for the census. The census experience of electronic data collection will also help Statistics NZ to better understand how this technology might improve other areas of statistical data gathering.

7. Where a development that is needed to meet a requirement does not have a clear fit within the core statistical functionality, we take a pragmatic approach and contract out. Whilst Statistics NZ have successfully developed a prototype online form and proved the feasibility of the approach, it does not have the financial resources or the internal skills necessary to develop a solution that would be capable of handling a projected peak of 31,000 concurrent users per hour for census night 2006, nor does the department want to develop these skills as a core competency. The development and operation of the online census for both the 2005 Dress Rehearsal and 2006 Census have therefore been contracted out under very strict controls to ensure the strict confidentiality provisions of the Statistics Act will be met.

CHALLENGES FACED AND LESSONS LEARNED

8. In developing the online census, a number of challenges were faced both from the business and technical perspectives. Whilst the technical perspective is extremely important, this paper focuses primarily on the business requirements, exploring how these have been met rather than the technical solution implemented. It only includes brief information from the Dress Rehearsal as we are still in the operational phase.

9. Broadly the challenges can be categorized as:

- managing the impact on existing business processes;
- managing Stakeholder Expectations;
- statistical; and,
- technological.

MANAGING THE IMPACT ON EXISTING BUSINESS PROCESSES

Integration into Field Enumeration

10. The most important element of the online census is successful integration into the field enumeration phase. Without this there could be no online option for the 2006 Census. This option increases the expectations of the public for immediate updates to field staff of forms submitted to census processing. Our aim has been to ensure that the respondent who completed their census form via the Internet one evening, is not faced the next day by the enumerator returning to pick up their paper form. The immediacy of the Internet heightens community expectations that the same use of technology permeates across the whole enumeration processes.

11. Feedback on the forms collected online is therefore communicated to the field using mobile phone text messages that are bundled and dispatched at different points in the day to coincide with field staff work patterns.

12. Prior to delivering forms in the field, the enumerators were given access to the online census facility to enable them to familiarize themselves with the online option, before it went live in the Dress Rehearsal. This was invaluable as an aid to their understanding of how the option would be viewed by the respondent prior to standing at their doorstep and offering that option.

13. Additional training for the field enumerators was also introduced to ensure they understood what is required of the respondent regardless of which option is the chosen method:

- firstly to ensure that the online option is communicated correctly to the respondent on the doorstep; and
- secondly, that the field communication systems integrating the online option with the paper option are correctly performed. This will have an impact on both the time taken to communicate the extra procedures and the amount of information that needs to be absorbed by the enumerators before they enter the field.

14. The training theme continues with the challenge facing field supervisors in respect of

being able to use the field operations monitoring systems. These are the mechanisms used for monitoring how the field responds to text messages they receive regarding the lodgement of Internet forms and related helpline actions. The complexity of these systems and the need to integrate with existing systems for the paper form enumeration process have posed a real challenge to ensure seamless operations, and it will be essential for the 2006 Census that this integration process performs as required.

Communication/Promotion

15. How the option is communicated to the population will have an impact on both who the eventual respondents are, as well as the level of uptake. Promotion of the online census to the desired audience will be key to managing the eventual level of uptake.

16. Consideration is being given to promoting the online census to particular "Internet savvy" groups and other sectors of society where we have had less success using the traditional collection methods e.g. apartments, Asian community (research has shown this sector of the community to be more receptive to completing online), students, and rural areas.

17. Complicating matters further, there will be two separate requirements to meet the needs of users for the Internet based solution for the 2006 Census:

- online census form completion; and,
- online general census information.

Online census form completion option

18. The first is to facilitate completion of the Census form via an "Internet Response Option" (IRO). The specifications call for a highly secure, simple to complete, stable online application that allows high volumes of Census forms to be completed over the limited census data collection period. Each form must be completed within a single session as there will be no capability to resume partially completed forms. The online census also has a concise bilingual context sensitive help facility included, the content of which has been based on the level of online support expected by the Internet respondents.

Online provision of general census information

19. The second requirement is to provide a comprehensive source of more general information relating to the census. This will provide information relating to the current census, along with assistance in completing the paper based census form. Information will include FAQs, multilingual help (for "paper based" respondents), request for forms and an ability for Statistics NZ to dynamically update certain areas of information on a regular basis. This information will be available for an extended period prior to and following the census event.

20. The online provision of general census information is consistent with research which has shown that accessing a website or using email are the most preferred sources of assistance. The 2003 Government Online survey (GO2003) identified that online information including email enquiries was preferred by 48% of New Zealanders with telephone being the next highest medium favoured by 30%.

21. The online census option needs to be completely 'locked-down' for security reasons, and only available for a relatively short space of time (6 weeks). In contrast, the more general range of general census information needs to be updateable on a daily basis and available for a much longer period of time (7 months). This creates an interesting dilemma in terms of how to promote and communicate the availability of each without compromising either.

Availability / Peak Activity Periods

22. The GO2003 survey found that 75% of New Zealanders have used the Internet in the month prior to the survey. New Zealand had the highest Internet usage amongst all participating countries and well above the global average of 45% (see Appendix A for participating countries).

23. The same survey identified that completing the census form online was perceived to be significantly useful by 69% of the New Zealand population. The highest sectors being the University educated (88%) and higher income earners (\$70-100K 84%; \$100K or more 89%). This will prove useful when considering the target audience for promoting the option.

24. In reality, testing has shown that actual uptake does not mirror the initial interest shown on the doorstep. Uptake is more dependant on the particular motivation of the respondent. Our own follow-up surveys have identified that the most common reason for using the online option is that there is no subsequent need to wait for an enumerator to call to pick up completed forms. From the results of our field tests in March and November 2003 we project that around 15-20% of people may take up this option in 2006. That's around 600,000 people or 250,000 households.

25. Recognizing that the online option for the 2006 Census could be New Zealand's largest online event to date, the system is being designed for stability and performance when there are high volumes of concurrent users including back up sites, multiple servers and substantial redundancy. In the event that there are still more concurrent users than the system can handle, a message will display for users advising them to try again later.

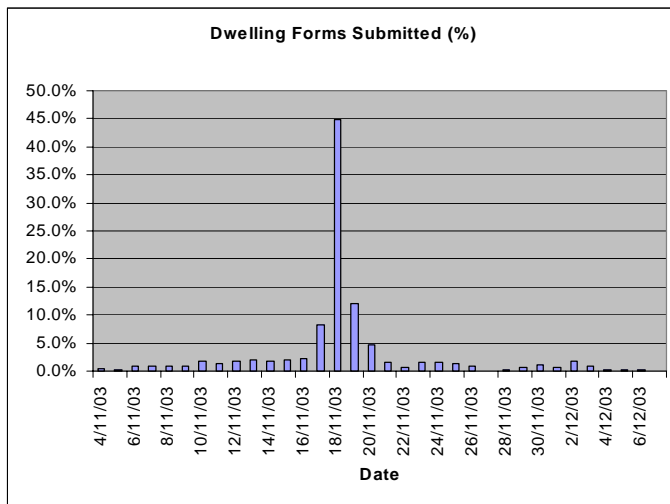
26. As part of our communications and risk management strategy, Statistics NZ is also talking to the telecommunications providers and ISPs (Internet Service Providers) to integrate network load planning for the 2006 Census.

Uptake

27. New Zealand's census is focused on a single day, with the prime activity-taking place on one night - known as 'census night'. Statistics NZ is designing the system for a spike of activity on census night and is confident that the system will handle this predicted load. Testing to date has shown that while a number of people complete their online census forms in advance of census night or on the day after, the majority complete their forms on census day, as required.

28. The chart below shows there was a spike of responses on the November Test day of 18 November 2003. A large proportion of the Dwelling Forms (DF) was received on the Test day (45%). This pattern has also been repeated during the recent Dress Rehearsal in March 2005.

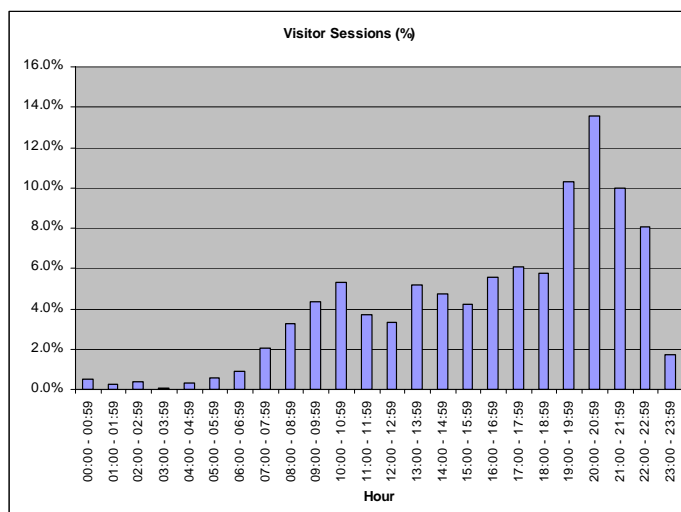
29. Adding to the overall challenge is the fact that the projected peak activity of Internet traffic will coincide with the heaviest demand for help over the public telephone helpline as witnessed by the activity for the 2001 Census. The competing demands for use of the nation's telecommunications infrastructure need to be well managed to mitigate the risk of a telecommunications 'meltdown'. To address this risk Statistics NZ has established joint risk workshops with all of the communications providers for the 2006 Census.



Usage by hour of day

30. The chart overleaf shows the website visitor sessions for 18 November 2003 by hour of day. Again, this pattern of usage across the day has been followed by the respondents in the March 2005 Dress Rehearsal.

31. The chart clearly identifies that most website visitor sessions were recorded from 1900 hours to 2300 hours with the peak hour between 2000 hours and 2100 hours. The system is being designed to cater for this expected peak website visitor session activity. Again, the competing demands of the public telephone helpline have to be considered in this narrow window of service demand. The demand profile for the helpline in the 2001 census shows that peak activity for the helpline was between 18:00 and 20:00. There is therefore an overlap from 19:00 to 20:00 where the telecommunications infrastructure would be handling a peak of both voice and data connections.



MANAGING STAKEHOLDER EXPECTATIONS

E-Government requirements

32. The E-Government strategy requires the following strategic alignment by 2004, 2007 and 2010 for all major government agencies including Statistics NZ:

- by June 2004, the Internet will be the dominant means of enabling ready access to government information and services;
- by June 2007, networks and Internet technologies will be integral to the delivery of government information and services; and
- by June 2010, the operation of government will have been transformed through its use of the Internet.

33. Statistics NZ is already strategically positioned to meet the 2004 requirement through its own web portal which provides ready access to a wealth of statistical information and services.

34. When the 2007 requirement is considered, there is an additional need to be looking at those phases of the survey cycle which are concerned with other interactions with the NZ public - whether in terms of interactive information provision, or in terms of online provision of data.

35. Meeting the requirement for 2010 will be the most challenging as this requires that services across Government are integrated seamlessly to ensure that the respondent does not have to provide the same information to one department that has already been provided to another. Meeting this requirement is not considered appropriate for the census as it would lose the independent and impartial position that it currently occupies. This could jeopardize the accuracy of the data supplied by individuals to census if it was felt that data supplied would be freely shared across government.

Security/Privacy and Confidentiality of Information

36. Public perceptions regarding security of the online census option will inevitably contribute to final Internet uptake levels. It has already been established that the current perception of the public to the Internet being a secure place to transmit data is not very high in New Zealand:

- an SNZ Public Attitudes Survey (carried out in April 2003) showed that only 5% (out of 700 respondents) felt that the Internet was the most secure method of data collection in comparison to the other methods available. Interestingly, those who were comfortable with using the Internet and generally used it frequently, believed it to be more secure;
- the GO2002 Survey commissioned by the e-Government unit (1000 respondents) found that only around one-third (31%) of New Zealanders consider it safe to use the Internet to provide the Government with personal information; and
- results from the GO2003 survey showed that only 35% of New Zealanders consider it either "Very Safe" or "Safe" to provide personal information to Government over the Internet.

37. Although public perception of the adequacy of Internet security seems to be rising, it would only take a single large negative event to reverse that trend. Any approach involving the security of data needs to be managed with these issues in mind. For instance, what would be the expected response from the public to find out that sensitive data they provide during census was stored on a private company's own server housed on their own premises, not government premises? And, if we promoted the Internet as the most secure environment to transmit data out of all the options available to respondents, what effect would that have on those people who did not have the option to use the Internet? Would they then mistrust the enumerator and consider that the information provided on the paper forms would not remain confidential? The right balance therefore needs to be found in the messages that are given to the public regarding use of the Internet as a secure environment.

38. Data provided by respondents over the Internet needs to be secure from a number of perspectives. The public already has an expectation, through online banking and shopping, that they will need to logon to the site and be authenticated and, that their transactions will be encrypted. Census needs to meet these expectations if not exceed them. The balance struck however between the perceptions of the public to the integrity of a secure logon and authentication process and how complex it is will again be an important factor in determining the level of uptake for this option. If it is believed to be too time-consuming then respondents may well abandon the online census in favour of the paper forms. Testing so far has shown that our logon process is not considered onerous.

39. Although data will be held externally to Statistics NZ when it is first captured it will be subject to stringent physical and logical controls. The data centre facilities will be physically and technically separated from all other services at the two hosting sites; under the full control of Statistics NZ personnel during both operation and testing; and, decommissioned fully following the 2006 Census. The collected data will be encrypted at all times and transferred to the Census Processing site via a secure virtual private network. All data collected external to Statistics NZ will be deleted from the data centres every 24 hours to ensure there is no possibility for unauthorized access.

STATISTICAL

Usability/Accessibility

40. On the other side of the security and performance risks, is the requirement to ensure all online census functions are not only easily accessible to users, but also easy to use. By making access to the online form too complex, there is a risk of both lower uptake and negative publicity. If respondents are familiar with using Internet applications and transactional websites such as online banking or shopping, they will find it relatively straightforward to complete their census forms online. Statistics NZ has tested the design internally and in public trials and usability has been key during the form's development.

41. The online census has been developed to represent, as closely as possible, the paper census forms in an online presence. Transferring a paper form to an electronic format is a process that whilst simple in theory is far more complex in practice. There are a number of processes for the paper form that are conducted manually by the field enumerators that would fall to the online respondent so increasing the burden on the respondent. In this sense, we have both increased and decreased the burden on the respondent.

Level of Respondent Burden	
Increase	Decrease
On the paper forms the enumerator performs a manual check that the six compulsory fields have been completed. For the online form completion, this is done at the end, as the respondent is asked to confirm the details provided, and is unable to submit the form if one or more of the compulsory fields have not been completed.	Based on responses to certain questions the respondent is 'routed' past non-required fields - following the strict intentions of the paper form without having to think about which question to answer next.
Field validations to remove the provision of 'impossible' values from respondents' responses. This gives rise to online checking of data being entered and could be considered to add to the burden of form completion as there is no such check during paper form completion.	Summary of forms completed together with receipt numbers is provided on the Household summary page helping the respondent to see which forms have been completed online and which forms remain to be completed.
Relative complexity of the logon process adds burden to accessing the online form.	By electing to provide census information through the online form, the respondent removes the necessity for the enumerator to return to collect the completed form.

Understanding Impact on Data

42. The use of an electronic form carries with it the potential to increase or decrease the quality of the data being received compared to that of the traditional paper method. This brings with it the risk that in so doing, the difference between the two sources would be so great as to render the electronic data unusable as a combined source of data. Time series data would also be affected immeasurably. In practice however, testing has shown that the sector of population choosing the online option would provide the same high quality level of data regardless of the medium they supplied it through.

43. The completion of census forms online can be performed by any one in the household. It may be that one person takes responsibility for the whole household and completes all forms that are due, or individuals may take responsibility for submitting their own census data. The younger generation may even want to be involved and complete their own individual forms. Although Internet form completion has been shown to be mainly homogeneous during testing, there is nothing to prevent some members of a household completing paper forms instead. The enumeration processes will need to accommodate these eventualities.

44. With more than one collection mode there is the opportunity for a percentage of respondents to duplicate their census form, deliberately or accidentally. There is also an increased possibility that duplicate forms may be submitted as respondent complete and submit forms prior to census night and then change their intentions for census night. Although there are checks to ensure that only one form per individual is actually processed for the census data, there is a likelihood that this will increase the pre-processing effort to identify and remove the redundant duplicate forms.

45. Although testing has not shown evidence of partial form completion or item non-response, there remains a risk that this may occur. The online option has been designed with this in mind, as the respondent is unaware of the compulsory fields until the end of the form has been reached, at which point there is a prompt for the respondent to check the details provided for the compulsory fields. No online form can be submitted if entries are not present in all of the compulsory fields.

TECHNOLOGICAL

46. Unsurprisingly, a number of technological challenges have been encountered while developing the online option, not least of which was determining and implementing the solution 18 months prior to the 2006 Census. The external environment is constantly changing as each new virus, security flaw or software bug is addressed, so a new and more persistent threat is identified to take its place.

47. The technological challenges faced relate not only to the environment that the online option is hosted within, but also to the technology used to build it. There is currently a mix of 3 computer operating systems and 4 Internet browsers that need to be tested for compatibility against the online option. As these are upgraded, patched or simply replaced, so the testing cycle starts once more. Simply, compatibility testing is a very time-consuming, but necessary

element of the development process to ensure the widest possible coverage for both the respondent and Statistics NZ.

48. The safety of the online option has been secured using a combination of the available technologies to achieve the required level of security including:

- Secure Socket Layer (SSL) with 128 bit encryption on the online census form pages;
- a firewall between: the Internet and the census web site; and, the census web site and the internal Statistics NZ integration servers;
- up-to-date anti-virus software on the servers;
- intruder detection and lock out facilities including excessive international traffic;
- two data centres hosting the web site in different locations to provide for disaster recovery, high load sharing capability during peak periods and the ability to take a server off-line for data deletion; and
- logon and authentication of users with the Household ID (11 char long) and ePIN (12 char long).

49. An interesting point to note regarding the logon and authentication process is that its main vulnerability, from a usability perspective, arises from the manual process of the enumerator writing in by hand, the last 4 characters of the Household ID onto the paper form. Tests have shown that this is where the respondent has the most trouble with the logon process due to the enumerators' handwriting either being misread, or the characters being incorrectly written. To overcome this, the helpline needs to have a checking process to ensure that the Household ID the respondent is trying to use is a valid one for authentication and entry to the online option.

CONCLUSION

50. The learning gained through both the use of this technology and impact on respondent behavior will prove invaluable as both government and public expectations increase for the 2011 Census. The expectation of government will be that future operations will be carried out using less resource, consequently at a lower cost. The public expectation will increase to the point where it will become an expectation as of right to be able to provide information to government through electronic means. It is essential therefore that Statistics NZ successfully implement the online census in 2006 and use the knowledge gained to plan for a more efficient and effective 2011 Census where real savings will be both expected and realized.

References: 2003 GO Survey – <http://www.e-government.govt.nz/docs/go-survey-2003/chapter1.html>

APPENDIX A

List of countries taking part in Government online 2003 (GO2003) survey:

Australia, Bulgaria, Canada, Czech Republic, Denmark, Estonia, Finland, Faroe Islands, France, Germany, Great Britain, Hong Kong, Hungary, India, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Malaysia, Netherlands New Zealand, Norway, Poland, Singapore, Slovakia, Spain, Taiwan, Turkey, USA.

* * * * *