Distr. GENERAL
12 September 2013
WP 13
ENGLISH ONLY

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection
(Geneva, Switzerland, 25-27 September 2013)

Topic (ii): Managing data collection functions in a changing environment

STATISTICS NEW ZEALAND NAVIGATES TOWARDS ONLINE DATA COLLECTION: PLANNING AND IMPLEMENTING AN INTERNET-BASED SURVEY

Working Paper

Prepared by Emma Mawby, Helen Smith and Colin Hewat, Statistics New Zealand

I. Introduction

1. This paper describes Statistics New Zealand's work to include an online option for business surveys. We discuss our reasons for introducing the new mode, the successes and challenges of designing this, the reasons we decided against going ahead with the first survey we identified as a candidate for this mode, and how this experience will influence our future decisions.

II. Background

2. Statistics NZ leads New Zealand's Official Statistics System. The department is governed by and administers the Statistics Act 1975. As New Zealand's major producer of official statistics, Statistics NZ is required to produce high quality, timely, and fit-for-purpose statistics.

3. Statistics NZ has already experienced success in implementing an Internet mode in the two most recent Census of Population and Dwellings (2006 and 2013). In 2006, we offered the Internet mode simultaneously with the paper mode and seven per cent of census forms were completed online. In the 2013 Census we collected over 5.6 million forms (both individual and dwelling forms) with just under two million of these completed online. This represents about 35% of all completed census forms, which was our target. During the 2013 Census a pilot was run for about 5,000 people in one geographical area of New Zealand. People in the pilot were offered the Internet mode initially, although they could subsequently request a paper form if they wanted one. While results are still being worked through for the pilot area, more than 65% of census forms were completed online.

4. In 2010, Statistics NZ proposed a 10-year programme of change to address several challenging issues, including rising operating costs, ageing IT systems, and outdated methodologies. The organization-wide programme of change, Statistics 2020 Te Kāpehu Whetū (Stats 2020), was launched in 2011. It aims to create a responsive, reliable, and sustainable organization by improving performance and safeguarding quality. Statistics NZ identified four strategic priorities to guide its work to 2020:

a. Lead the Official Statistics System so that it efficiently produces the information that New Zealand needs.

b. Obtain more value from the country's investment in official statistics.

c. Transform the way Statistics NZ delivers statistics.

d. Be a more responsive, customer-focused, influential, and sustainable organization.

5. In its first year, Stats 2020 focused on stabilising IT systems, developing new operating platforms for analysing data, and standardising the production of statistical outputs. In its second year, the Transform Collections Programme built momentum.

III. The Transform Collections Programme

6. Transform Collections is the programme of work that supports Statistics NZ's third strategic priority: transform the way we deliver statistics. The programme aims to design, develop, and implement a data collection platform or future operating model that will harness current technologies. This model will help Statistics NZ achieve its vision of being a world-class collector and producer of statistics, in a way that is efficient, cost effective, and responsive to customers and respondents.

7. Statistics NZ recognises that the environment in which it operates is complex. People are more technologically capable, using the Internet for both business and social networking. Government is expected to be equally technically capable. To meet the needs and lifestyles of its customers and to remain efficient and flexible, the organization needs to collect and disseminate information using modern technologies that can adapt to change.

8. One of the major objectives of the Transform Collections Programme is to put in place the necessary changes to business processes and supporting tools (e.g. guidance, forms, templates, notices), and the existing IT system. These changes will allow us to collect business surveys online. This is no small task, as it requires balancing our aims of reducing collection costs and improving respondent experience. Respondents must be confident that we are trying to reduce the load on them. Users of our data must be assured that we are delivering good-quality, relevant outputs and that we are able to respond to their changing needs.

9. This transformation work required a marriage of skills from different areas: collection staff, questionnaire designers, statistical methodologists, subject matter experts, IT personnel, and project managers. Team members had various levels of experience at Statistics NZ. Additionally, a balance needed to be struck between the need for a timely and practical solution and methodological concerns that would require time to explore and address.

IV. The Interim Internet Project

10. Statistics NZ's experience from previous change programmes shows that a staged and sequenced approach is the most appropriate for developing new modes and platforms. First, we decided to focus on transforming business surveys using existing databases and systems. One of the short-term objectives of the Transform Collections Programme is to introduce a basic Internet mode to business survey respondents. The corresponding long-term objective is to provide a more comprehensive Internet mode for both business and social surveys.

11. The project started with defining the solution scope, setting priorities and defining key success criteria. Stakeholders responsible for data collection, questionnaire design, survey design, and production of outputs were involved in this process. These criteria included the ease of migration, logon and authentication processes, respondent specific pre-population of questionnaire fields, and customisable survey communications.

12. Next, we identified potential issues associated with Internet surveys, and then researched reliable solutions, new possibilities, and current practices from around the world. We consulted with the Australian Bureau of Statistics about their experiences with implementing an online mode in their Monthly Population Survey. This research enabled us to identify some of the challenges that occur frequently: availability of sampling frames with good coverage of email addresses; mode effects; privacy; security; authorisation; and authentication. One of the greatest challenges from a methodological perspective was whether we should accept the presence of mode effects and attempt to measure them, or if we should try to minimise them at the questionnaire-design stage of the survey cycle. After investigating both possibilities we chose to minimise mode effects, as discussed below.

V. Challenges

A. Method of first contact

13. Initially, the idea of a totally paperless collection seemed very appealing. We would email questionnaires to respondents who would return them the same way. Respondents could get help to complete their questionnaires by checking frequently asked questions about the survey on our website or contacting the survey's helpdesk. Unfortunately, the quality of the coverage of email addresses on the Statistics NZ Business Register does not yet match that of postal addresses, so using email as a method of first contact could result in coverage error. Given this, we decided to make initial contact post or telephone. This would be an opportunity to collect email addresses that could be used for future contact.

B. Mode effects

14. There is much literature that states that the same question asked through different modes (eg by telephone, face to face, Internet, or paper) can elicit different responses from the same individual, or that certain types of errors are particular to certain modes. These are called 'mode effects'. We wanted to closely manage the risk of mode effects as we thought the impact on data quality would be significant. One way of investigating the size of mode effects would be to double the sample size, randomly assign businesses to each different mode, and then compare the results. However, this was rejected due to the high cost and increased respondent load resulting from a larger sample size. Instead we chose a survey with a low risk of mode effects (based on the type of questions and the variability inherent in the population) to be the prototype Internet survey and to minimise mode effects by using questionnaire-design techniques.

15. We also had to decide how to manage modes. Best practice principles identified in international literature state that multiple modes should be offered to respondents sequentially (starting with the cheapest option – which is the Internet – then moving to paper, and finally following up with telephone interviews). This method is the most cost-effective solution because it maximises the number of Internet responses, and is simpler to analyse than if all modes were offered at the same time. Offering all respondents a choice of all available modes was also prohibited by cost.

C. Respondent experience

16. Offering all modes at the same time could provide a better respondent experience, because respondents can respond in their preferred mode. However, work by Dilman, Millar, and O'Neill (2009) also shows that respondent preference can be influenced by the choice of modes they are presented with. In particular, once a respondent tries a non-preferred mode of response (e.g. the Internet) their preference for that mode is expected to increase until it becomes their preference. This effect would benefit all future surveys and align well with the programme objectives; therefore, the decision was made to offer modes sequentially. An Internet response mode offers Statistics NZ the opportunity to reduce both cost and respondent burden and to improve the quality of responses. Statistics NZ assesses and manages respondent burden for surveys based on the mean time taken to complete them (respondent load).

D. Privacy, security, authorisation, and authentication

17. New Zealand does not currently have a system of unique business numbers[1], so it is not possible to authenticate the identity of businesses in the same way we authenticate individuals. In the paper mode collection, authentication and authorisation are dealt with by pre-filling the name and address fields of the questionnaire. The questionnaires are then mailed out to the person (or job title) responsible for completing the questionnaire, who returns it by prepaid mail. Details of the person (or job title) responsible for completing the questionnaire are recorded and maintained by the subject matter area. Statistics NZ assigns a relationship manager to businesses known to make a large contribution to the New Zealand economy (and that are selected for several business surveys each year) to ensure good communication.

18. An all-of-government initiative to provide an identity authentication service focuses on resolving the privacy and security issues faced by users to accessing government services using the Internet. Because the service is an all-of-government initiative, Statistics NZ is committed to implementing an Internet mode using its functionality.

19. This system currently authenticates an individual's identity securely when they access government services online. However, it cannot assess whether an individual is authorised to respond for a business. Additionally, it was difficult to measure how many businesses in any survey population had registered for the service, as there is currently no government obligation for people to do so. We were aware that Internet response rates may be compromised if the people responding for the businesses were not familiar with the service.

E. IT requirements

20. The most difficult aspects of the questionnaire design was testing the front end of the IT infrastructure (the website landing page and logon processes) and ensuring the questionnaire connected with the back end of the IT infrastructure (i.e. that data was displayed, saved, and processed as expected).

21. Construction of IT infrastructure and the establishment of the servers took longer than planned. Without the full IT system running, we could not test these issues, which made us doubt whether the elements in the system would work together successfully. With no full IT infrastructure in place, there was a higher risk of respondents not being able to complete the Internet questionnaire.

VI. Selecting the questionnaire design tool

22. We began the tool selection process with a thorough requirements-gathering exercise. Questionnaire designers and others with knowledge of the functionality needed provided a list of detailed and specific user needs. We documented these requirements and used them as a measure against which to score potential tools.

23. We did not consider building a tool in-house for several reasons: an all-of-government policy that encourages buying commercial, 'off-the-shelf' tools; limited time; constraints of the project; and cost. Instead we focused on assessing a shortlist of potential 'off the shelf' tools. Only a few tools met our security requirement of having a cloud source based in New Zealand, and being within the budget of the project.

24. We demanded much functionality from the tool, such as the ability to manage complex routing, calculations, data piping, and validation. We also should be able to control the look and feel of the questionnaire to make it user friendly, so that questionnaire designers would not need IT support. During the assessment phase we had to make some compromises as none of the 'off the shelf' tools

[1] New Zealand is currently implementing a unique business number due in 2014 and is in the process of addressing the issue of authenticating individuals within a business context.

could fully satisfy all our requirements (e.g. having a save-and-resume function, or being able to execute complex calculations). The tool we chose satisfied the most number of requirements. It gave us a high level of control over the look and feel of the questionnaire and received positive reviews from users.

VII. Selecting the initial survey

25. Alongside tool selection, we selected a business survey to be the prototype Internet survey. We chose to introduce an Internet mode by carrying out a 'phased analysis, phased rollout' solution. This means that business surveys with similar business requirements were grouped together. These surveys would then have an Internet mode implemented and lessons would be learned from the initial implementation. The cycle would be repeated for further groups of surveys until all business surveys offered an Internet mode. This approach became known as the 'do-learn-do' approach.

26. To select the first survey, we evaluated factors such as complexity, sample size, organizational risk, range of question styles (i.e. quantitative versus qualitative, objective versus subjective), and risk of mode effects. We chose the New Zealand Energy Use Survey (NZEUS) as the first business survey to offer an Internet mode. NZEUS covers businesses in all industries in a three-year period. All industry estimates were first published in 2011. The survey collects information on the amount of energy used by businesses and categorical information about energy management practices.

VIII. Designing the questionnaire

27. We designed the NZEUS Internet questionnaire to maximise the benefits of the Internet, whilst retaining consistency with the existing paper questionnaire, which was being maintained for respondents who would not use the online mode. Through agreement with the subject matter area, we made decisions about which dynamic elements should be built into the form (e.g. routing, hiding sections, or automatic calculations), and what validation and edits should be programmed in (e.g. checking dates and email formats, data ranges, and selecting mandatory questions).

28. Building an Internet questionnaire was straightforward once we had mastered the software. We successfully built dynamic elements into the form and the questionnaire design tool proved it could cope with complex routing and varied situations. The external user interface largely met our expectations and a good look and feel for the form was achieved. We conducted a first phase of usability testing and the Internet form received positive feedback from respondents, especially about the visual aspects and the ease of use of the questionnaire.

IX. Implementing the Internet mode

29. The timeframes for delivering the Internet mode were ambitious due to the timing of the release of NZEUS. Once we decided to offer the Internet mode sequentially, we had to decide whether we should make the Internet mode available to all the businesses in the sample or restrict it to a small proportion. While we wanted to implement solutions that were sustainable in the long-term, the project also offered opportunities for research. We needed to balance our desire to learn about mode effects with our need to produce estimates of equal or better quality as the previous NZEUS collections.

30. Our final decision was to first offer the Internet mode to all businesses in the sample. Businesses could ask for a paper form any time during the collection period. We would use milestones during the collection period to monitor the response rate (as we do in paper collections). If the response rate did not meet our expectations at a critical milestone we would send all non-respondents a paper questionnaire to try to increase the response rate.

31. Limited development time meant we had to remove some requirements for the first phase of work. These were to be delivered later in the project. One of these requirements was the ability to pre-fill

fields with existing data, such as business address details. To add this functionality, we needed more time to ensure acceptable standards of security and privacy around prepopulating fields were met. Unfortunately, we did not fully realise the implications of this until late in the development of the solution. At this point we were concerned that businesses with complex structures, or respondents completing forms for multiple businesses, might make errors if the business details were not pre-filled. This issue is a particular problem for NZEUS because it can be difficult for businesses with complex structures to allot their energy use to the collection unit required by Statistics NZ. Also, this meant that the Internet form needed more respondent input than the paper form, with a potential adverse effect on respondent experience.

32. At the same time, an issue arose about using the all-of-government shared authentication service given that it is designed to authenticate individual identity, not the identity of people associated with a business. During cognitive testing of the NZEUS questionnaire we asked a small number of respondents if they were aware of the service and if they could use this for a Statistics NZ survey. Feedback was mixed. Large businesses were unsure about using the service, but smaller businesses (run by individuals) saw no problems. The larger businesses were unsure if there was a business-wide account already set up, if they would have access to it if one did exist, or if they would be allowed to use their own individual account to complete the survey. These issues had the potential to pose a barrier to respondents completing a survey online and so we noted them. Unfortunately, the logon service is also undergoing change, so we were not able to resolve our questions in the time available, and unable to proceed with implementation of an internet mode for the NZEUS collection.

X. Lessons learned

33. Twelve months into the Transform Collections programme, we are evaluating our progress, recognising our successes, and identifying areas that need further development. By offering an Internet mode for our business surveys we are modernising our systems, improving how government interacts with businesses online, and providing an improved customer experience.

34. We learnt that what we thought would be our greatest challenges – such as mode effects and method of first contact – can be dwarfed by unexpected and sometimes show-stopping issues, such as pre-fill security and the suitability of the logon process. We realised we should not assume anything, from the seemingly obvious requirement of identifying the mode of response, to the expectation that government tools will effectively meet our needs. Additionally, the quick pace of change in the external technology environment means it is difficult to implement an interim solution.

35. We learned many lessons on this journey and are in an excellent position to provide an Internet mode for business surveys in the future. We remain committed to providing an Internet mode in business surveys to ensure that New Zealanders can complete their transactions with government easily in a digital environment. Our greatest insight is to understand clearly why a survey is being moved to the Internet, and then use this knowledge to create unified objectives. If the primary objective is to reduce costs and save money, then we will use a different approach than if the objective is to reduce respondent burden. If the objective is to increase response rates from a specific sub-population, then the approach will be different again. Statistics NZ has an obligation to implement the all-of-government initiative to improve interaction with government. However, we need to clearly understand how this initiative fits with our aims as a producer of statistics. We must also recognise that although we can learn much from solutions implemented by other agencies, our biggest challenges may be unique to our own situation.

XI. References

Millar, M, O'Neill, A C, & Dillman, D. (2009). Are mode preferences real? Technical Report 09-003. Washington: Social & Economic Sciences Research Center.