

The Role and Challenges for Official Statistics in Policy Evaluation

Paper presented at the Australasian Evaluation Society conference,
at Canberra, 4 September 2009

Geoff Bascand

Government Statistician, Statistics New Zealand

Abstract

Official statistics are a vital part of the policy and programme evaluation process. Statistics New Zealand works closely with policy agencies and researchers to create new, policy relevant sources of information and extract value from existing datasets. This paper looks at the environment that Statistics NZ operates in and how this relates to policy evaluation and the increasing options for micro data access that Statistics NZ provides. Examples are given of innovative datasets that have been built in collaboration with policy agencies. The paper concludes with a discussion of benefits and challenges that arise, including improvements to the policy relevance of our official statistics, boundary issues that confront official statisticians and hence evaluators around confidentiality and statistical release practices, and challenges of integrating programme intervention information with statistical information.

A full version of this paper is available from Statistics New Zealand.

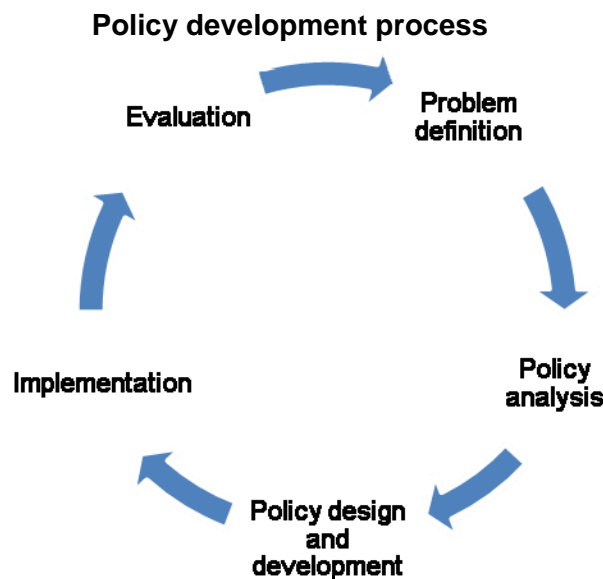
Acknowledgement

I thank Tas Papadopoulos in particular for assistance with this paper.

1. Data as evidence for evaluation

Official statistics play an integral role in the policy development process illustrated in figure 1 (State Services Commission, 1999).

Figure 1



Policymakers need data to help at least four stages of this cycle.

First, high level data at a high level of aggregation can help them define or identify a problem – national output is lower than that of our contemporaries or is growing at a slower rate.

Secondly, data is evidence of how the economy and society is functioning (or malfunctioning) – disaggregated socioeconomic data can identify educational, ethnic, or family developmental factors. This provides a base for the analysis that leads into policy design.

Thirdly, they need data to help policymakers understand the process of implementing and using the policy – what is delivered and to whom?

Finally, data can help evaluate the impact of policy – the ultimate measure of the effectiveness of policy is the impact on the high-level outcomes such as gross domestic product (GDP) or unemployment.

Microdata gives policymakers the evidence to distinguish policy and programme impacts from other influences, to assess marginal changes for more or less affected groups or periods, and to identify side-effects (positive and negative). This evaluation allows policymakers to:

- see if the benefits of a policy or intervention exceed the costs in terms of taxes raised to finance them and other consequences (foreseen or otherwise)
- determine whether the policy should be continued or removed
- know how they might improve the policy/regulation or create a better one
- understand how the economy, society, and environment function.

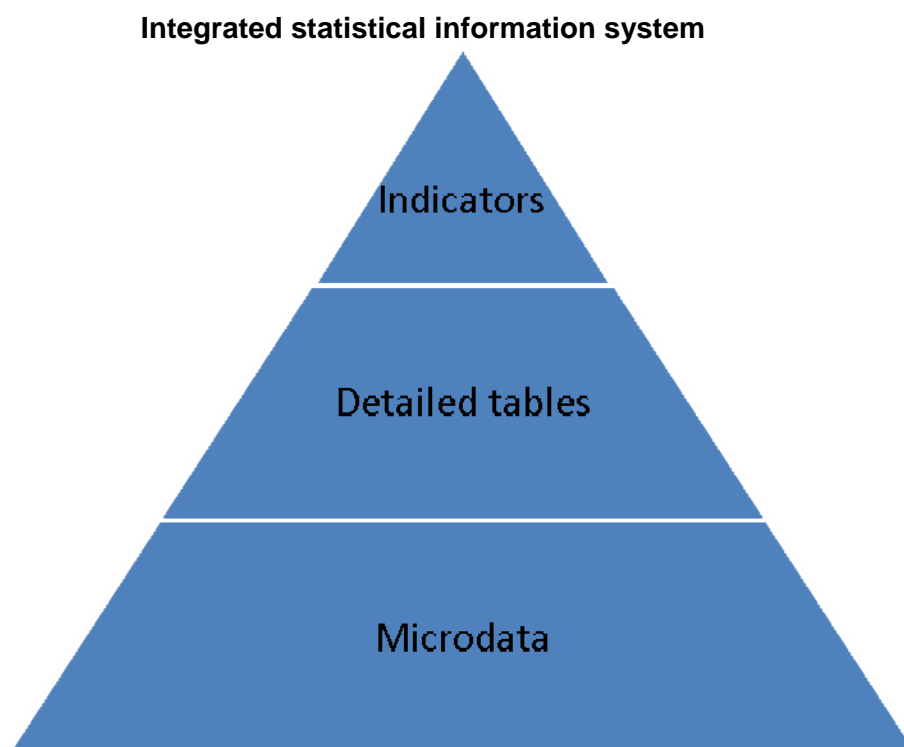
The demand for appropriate data seems obvious, yet is often hard to fulfil. Banks (2009) said that after methodology, his second essential ingredient for evidence-based policy was data. However “a major failing of governments in Australia, and probably worldwide, has been in not generating the data needed to evaluate their own programs” (p12), and he expressed concern that funding for data collections is being cut.

Sadly, two reviews of evaluation practice in New Zealand (Lunt & Trotman, 2005, and Ryan, 2003) appear to take data availability for granted. While exhorting strategies to grow evaluation capabilities and examining the challenges in generating an effectiveness orientation informed by research and evaluation, neither mentions the relevance and availability of data at all.

2. Official statistics and policy evaluation

While this paper concentrates on how microdata can help policymakers gather evidence and evaluate policies, official statistics can be useful in a range of ways and formats. Figure 2, adapted from one developed by Dennis Trewin (2008) represents an integrated statistical information system delineated by level of detail.

Figure 2



Taking an example, the national unemployment rate from a labour force survey is an indicator, which is supported by detailed tables of labour force status broken down by various demographic characteristics. The actual individual-level labour force survey dataset is the microdata.

Traditionally, detailed tabular statistics have been the basis of much policy assessment. Detailed time series and industry sector tabular statistics have been used extensively in economic and social research and evaluation. For example, evaluation of the impact of New Zealand's Working for Families tax-credit assistance programme on sole parents' employment status used tables of the Household Labour Force Survey (HLFS), and sectoral modelling analyses of tax and regulatory policies, and recent analyses of the impacts of climate change policy responses have been undertaken in this way.

However, these types of analysis may lack the power to identify the effects of different policies or other causal factors. For this, it is often necessary to estimate the effect of policy changes on individual agents. As technology and techniques to do this have improved, so the demand for microdata, longitudinal if possible, has increased.

The rest of this paper looks specifically at uses of microdata in policy evaluations.

3. International principles

National statistical offices (NSOs) need to balance impartiality, independence, and confidentiality requirements with the need to produce policy-relevant statistics in policy research and evaluation (United Nations, 2007, and United Nations, in press).

Statistics legislation, official statistics principles, and the standing and role of the national (government) statistician are critical for managing the tensions that can arise. Despite the challenges with achieving this balance, NSOs need to move more towards producing statistics that are more relevant and accessible, and so support policy advisers and researchers.

There are strict protocols around what access a national statistical office can give to microdata. The 57th Conference of European Statisticians released new principles on statistical data integration. Principle 1 states that “data integration should only be undertaken by NSOs for statistical and research purposes”.

4. New Zealand’s legal environment

In New Zealand, access to microdata is governed through three Acts.

The Statistics Act 1975 covers how Statistics NZ collects, uses, securely stores, and manages access to data. Data obtained under the Act is to be used only for statistical purposes. Research is not a defined purpose under the Act, and interpretations have varied as to whether it is encompassed within the meaning of ‘statistical purposes’ or not. Usually it has been regarded as an insufficient purpose in its own right.

The Act provides the Government Statistician with the discretion to approve data access by other government departments for statistical or research purposes under tightly controlled conditions, and to employ independent contractors to carry out work that enables Statistics NZ to deliver quality official statistics.

The Tax Administration Act 1994 allows Inland Revenue to provide data to Statistics NZ under the discretion of the Commissioner, who needs to be satisfied that the integrity of the tax system is not going to be compromised by Statistics NZ’s use of the data.

Statistics NZ is a trusted organisation with the systems in place to manage access to microdata for evaluation and research in a number of ways. Trewin (2008) notes there is room for improvement, and that although access to Statistics NZ microdata is generous for government department researchers, there are significant restrictions for other researchers. These restrictions are due to the wording of the Statistics Act, rather than Statistics NZ’s policies.

The Privacy Act 1993 protects the privacy of individuals. While there is no specific legislative requirement around data matching for statistical purposes, Statistics NZ is concerned that the privacy of individuals is safeguarded, and is seen to be protected.

Principle 1 of Statistics NZ’s *Data Integration Manual* states that “Statistics NZ must only undertake data integration if integration will produce or improve official statistics”. This is a narrower definition than that agreed to by the 57th Conference of European Statisticians. Principle 1 may limit the potential for programme administration data to be integrated with existing Statistics NZ datasets for research purposes.

5. Analysing and evaluating policy using Statistics NZ data

This section looks at examples of how policy has been analysed and evaluated using new, innovative, datasets built by Statistics NZ in collaboration with policy agencies. It

discusses the gains that have arisen from these datasets, as well as the challenges they offer for the providers of the statistics, and the end user.

6.1 Immigration policy and settlement services

The Longitudinal Immigration Survey: New Zealand (LisNZ) is conducted in partnership between Statistics NZ and the New Zealand Department of Labour (DoL), and provides data, not elsewhere available, to help:

- identify the outcomes of immigration policies
- generalise migrants' initial settlement experiences in New Zealand.

Migrants have been interviewed over three waves (beginning in 2005) approximately 6 months, 18 months, and 36 months after arriving in New Zealand.

The main advantage for policymakers when using this data is that the questionnaire has been designed to address the stated policy-related information gaps. Therefore, the DoL can easily apply the results to immigration policy developments and the design of migrant settlement services. A key preliminary finding is that those in the 'skilled principal applicant' category are more likely than other migrant groups to be working at the same or higher skill level as previously (which contradicts the popular impression of skilled migrants working in positions unrelated to their previous occupation).

The main challenges for Statistics NZ have been around the changing policy parameters that delayed the survey's introduction and the management of the data provided to DoL.

The Department of Labour who led the analysis for the first release of wave 1 data found the deadlines and restrictions on who could access the data difficult to work with, which held up the data release.

For the first release of wave 2 data, DoL handed more of the responsibility for data analysis to Statistics NZ. This has enabled their staff to spend more time on policy analysis. This includes releasing information from wave 2 progressively over a longer period. This should sustain interest in the survey between waves and reduce the quantity of information released in reports for future waves.

6.2 Income, saving, and family dynamics

The Survey of Family, Income and Employment (SoFIE) covers major influences on income, including employment and education experiences, household and family status and changes, demographic factors, and health status. SoFIE began in 2002 and will run for eight years or waves. Respondents are revisited yearly to build a picture of how their circumstances change over time.

The Treasury has used SoFIE to investigate whether New Zealanders are saving enough for their retirement. Results show large parts of the population aged 45–64 years have made adequate provision for their retirement, but about one-third may not be saving enough.

Questions specific to KiwiSaver (a subsidised, employer-based, voluntary retirement savings scheme, operational from July 2007) have been incorporated into the SoFIE questionnaire for wave 8. KiwiSaver aims to improve New Zealand's low average rate of saving, and is being evaluated by a group of government agencies. As SoFIE also collects information about people's superannuation and other saving schemes, the data will be used to evaluate changes in savings behaviour.

There are extra confidentiality risks associated with providing longitudinal microdata for researchers. As a result, Statistics NZ has a policy that it can only be accessed through our secure on-site Data Laboratory, which could be a barrier to using data for policymaking where timeliness can be a factor.

6.3 Life and death

The New Zealand Census-Mortality Study, a collaboration between the Ministry of Health, the University of Otago, and Statistics NZ (University of Otago, nd), links administrative death registration records with census records to measure socio-economic gradients in mortality.

The study has linked mortality records to census records for the three years following the censuses for 1981, 1986, 1991, 1996, and 2001, and offers several advantages over other study designs including:

- direct linkage of exposures to outcomes at the individual level
- rich exposure data
- great statistical power (because of the large coverage of the population).

The integrated census mortality database has resulted in over 50 studies addressing questions such as:

- What is the relationship between individual socio-economic factors and mortality in New Zealand?
- Is unemployment associated with suicide?
- What is the contribution of smoking to health inequalities?

In addition, the integration revealed the systematic misclassification of ethnicity information on mortality records. This meant that mortality rates for the Māori and Pacific populations had been underestimated.

A major issue with this project was whether the integration of census records would adversely affect public confidence in Statistics NZ. Any activity with the potential to erode public trust in the census is approached cautiously.

Statistics NZ decided that the benefits of the data integration for health research and policy outweighed the risks. The risks were managed by being open and transparent about the linkage of records, by not using direct identifiers in the integration process, and by restricting access to approved researchers within the Data Laboratory.

The results of the New Zealand Census-Mortality Study have had relatively wide exposure in the media. There have been only isolated cases of negative reaction to the linking, suggesting at worst public apathy and at best public acceptance.

6.4 Employment and income outcomes

Many programmes aim to improve employment or income outcomes. Evaluations have traditionally used administrative records, HLFS tabulations, and one-off surveys. Weaknesses in these sources are coverage (cell numbers of these affected by minimum wages are small in a sample survey), the difficulty of repeating ad hoc surveys, and that administrative data informs only about the period in receipt of service, not its subsequent impacts.

The Linked Employer-Employee Dataset (LEED) improves the information about labour market dynamics by examining the interaction of workers and firms by linking firms and workers both longitudinally and to each other. Data comes from monthly PAYE records for all individuals employed and includes all income types taxed at source, including income-tested benefits, New Zealand Superannuation, student allowances, and accident compensation payments. Annual self-employment tax returns are also integrated into LEED. All this information is integrated with the Business Frame, which is a register of private and public sector organisations in New Zealand.

LEED has a unique feature that makes it useful for policy evaluation. It can be used to assess post-intervention outcomes and can potentially permit control-group matching because it is a census (although demographic variables are limited). It is often difficult to evaluate outcomes once clients have left the reach of the administrative system or service provider.

A number of research initiatives have been undertaken using LEED, including:

- **outcomes of injury**, which estimated the effect of injuries on employment and total income by comparing the observed changes in outcomes for the injured population, with a matched 'control' group of non-injured individuals who have similar pre-injury observed characteristics
- **outcomes of benefit receipt**, which examined benefit-to-work transitions, which were not previously able to be examined using administrative benefit records.
- **outcomes of tertiary education**, a feasibility study on integrating tertiary education data (including industry training) with LEED. This will allow the labour market outcomes of educational achievement to be estimated.

Challenges that arose while LEED was developed included:

- understanding legal and relationship issues to do with accessing other datasets. Source agencies generally recognised the benefits; however, some were reluctant to allow access to politically sensitive data. This challenge was overcome through a slow process of building trust.
- developing statistical tools and information systems to integrate these large complex datasets and turn them into information. Capability issues were overcome by contracting in data warehousing expertise and seconding government researchers to do policy-relevant research.

Challenges that continue are the:

- need for careful monitoring of the security of the dataset and the confidentiality of any outputs, and the need to address the potential concern of individuals that unrelated information might be collected about them in an ever-growing database. Transparency has been a feature of the LEED project.
- lack of demographic information such as ethnicity and family structure.

While LEED is good for analysing the outcomes of groups of people being treated, it is often difficult to identify those who could have been treated under a particular programme, but were not. The integration of programme and demographic data might go some way to meeting this challenge. Current legislation and policy may preclude such an integration as it would need to be established that it benefits official statistics.

6.5 Business performance and the impacts of business assistance

Traditionally, industry analyses and supporting data disaggregations have informed much microeconomic policy-making. An increasing trend in New Zealand and worldwide is to analyse performance and productivity drivers at the firm level.

The prototype Longitudinal Business Database (LBD) is a longitudinal database containing business-related data from 2000 to 2007. It is constructed by linking a longitudinal version of the Business Frame with business survey and administrative data.

The LBD is used by researchers seconded Statistics NZ and working in the Data Laboratory, to analyse business practices, performance, productivity, hedging behaviour, merchandise trade, and finance.

The LBD provides a unique opportunity to evaluate government's business assistance schemes, through comparing the performance of assisted firms to unassisted firms. The Ministry of Economic Development's 2009 evaluation of the Growth Services Range, designed to accelerate the development of high-growth firms, showed that the programme had a positive impact on business growth that was greater than the cost of the programme. The authors contrast this with a "previous evaluation [that] was unable to provide conclusive evidence of the programme's impact on firm growth due to a lack of a control group" (Ministry of Economic Development, 2009).

7. Lessons from official statistics systems for evaluation practice

Official statistics systems have much to offer the evaluation community in terms of practices and policies. The New Zealand official statistics system has three great strengths:

1. Critical national statistics are designed and constructed by an independent agency with no responsibilities for policy-setting. The statistics are presented and released independently of the input of those with policy responsibilities.
2. Each statistical output is released with a comprehensive set of metadata about the methods used in its production.

These strengths relate back to Gary Banks' essential ingredients for evidence-based policy. A further two ingredients he cited are independence and transparency. Judgement in evaluation is unavoidable. Therefore, as he states (2009) "... evidence is more likely to be robust and seen to be so if it [is] not subjected to influence or barrow pushing". As to transparency of data and methods, "... no evidence is immutable. If it hasn't been tested, or contested, we can't really call it 'evidence'".

How well placed is government-based evaluation, in terms of achieving the required public confidence in the transparency and independence of reporting? The evaluation profession has strong ethics and professional practices supporting the integrity of their work.

The main lesson the profession can take from the official statistics systems is to ask: how well does the current structure and context for evaluations operate in regard to the:

- **design of the evaluation?**
How does the system deal with situations where the scope of an evaluation could be influenced to accommodate the needs of politicians or the accountable departments? How can evaluators demonstrate to the public that such direction is not possible?

- **release of the evaluation?**

Is the presentation and timing of the first public release of evaluation conclusions to the public done in a manner that is sufficiently and demonstrably transparent and independent so that it is not influenced by policymakers, providers or politicians?

8. Evaluating the evidence of the usefulness of official statistics for policymakers

Over the past decade, policy and programme evaluation has been significantly advantaged by new, policy-relevant, official statistics. In building these datasets, Statistics NZ has benefited from collaboration with policy agencies. Their sponsorship through the budgetary process was instrumental in establishing SOFIE, the LisNZ, and LEED, while data quality has benefited from the insights of users. Data will only improve if we sustain strong partnerships between data users and providers, especially in an economic environment where value for money is closely scrutinised.

The potential impact of evaluative research on policy design and operation is extraordinary. New Zealand grants around 50,000 permanent and long-term migrant approvals each year (or over 1 percent of the population) and, immigration has a significant fiscal impact. Government spending on social welfare and employment assistance exceeded \$23 billion (or 13 percent of GDP) in the 2009 tax year, affecting some 800,000 people and their families. Tertiary education spending amounted to 2.2 percent of GDP in 2009.

Policy evaluation and research that informs better designed policy and results in improved policy outcomes is crucial for enhancing New Zealander's well-being.

NSOs need to move towards more effective collaboration with policymakers and researchers, while observing institutional, technological, and methodological practices that maintain trust and confidence. These relationships are governed by the legal environment and official statistics principles. For example restrictions on access to Statistics NZ microdata can only be lessened by making changes to the Statistics Act – something that must be undertaken with great care and consideration.

Another issue for the New Zealand policy community is how to make the best use of the new person- and business-level administrative datasets that we have constructed. Hyslop (2009) has stated that LEED and the LBD have the potential to be the backbone of programme evaluation if programme data could be integrated. This has already been done with the LBD, as statistics on business growth could be generated, justifying the integration, and privacy issues around integrating firm level and programme information do not apply. However, integrating programme intervention information on individuals with LEED under the current legal environment and data integration principles poses a challenge.

The issue is how far statistical purposes should extend? At what point does evaluation become too policy-specific and the data use loses its wider public use aspect underpinning its compulsory acquisition? Work needs to be done to clarify where the integration of programme-related information fits in this continuum.

The continued development of New Zealand's new official statistics system can only lead to improvements in the quality of data for programme evaluation and policy-related research. It also has the potential to improve access to this information to facilitate more independent and transparent policy-related research in the future and better match research capability with data sources.

References

Banks, G (2009, February). *Evidence-based policy-making: What is it? How do we get it?* Speech to the Australian and New Zealand School of Government (ANZSOG)/Australian National University (ANU) Lecture Series. Available from:
www.pc.gov.au/speeches/cs20090204

Hyslop, D (2009, July). *Programme evaluation*. Keynote address at the New Zealand Association of Economists Conference, Wellington.

Lunt, NT, & Trotman, IG (2005, September). A stagecraft of New Zealand evaluation. *Evaluation Journal of Australasia*, 5(1), 3–10.

Ministry of Economic Development (2009). *Evaluation of the growth services range: Statistical analysis using firm based performance data*. Available from:
www.med.govt.nz/templates/MultipageDocumentTOC_40617.aspx

Ryan, B. (2003, August). Death by evaluation? Reflections on monitoring and evaluation in Australian and New Zealand. *Evaluation Journal of Australasia*, 3(1), 3.

State Services Commission (1999). *Essential ingredients: Improving the quality of policy advice* (Occasional Papers, 9). Wellington: Author. Available from:
www.ssc.govt.nz/display/document.asp?DocID=2910

Statistics NZ (2006). *Data Integration Manual*. Wellington: Author. Available from:
www.stats.govt.nz/about_us/policies-and-guidelines/data-integration.aspx

Trewin, D (2008). *The future role of a national statistical office*. Unpublished paper. Available from Statistics NZ.

United Nations (2007). *Guidelines and core principles for managing statistical confidentiality and microdata access*. New York: Author. Available from:
www.unece.org/stats/publications/Managing.statistical.confidentiality.and.microdata.access.pdf

United Nations (forthcoming 2009). *Report of the fifty-seventh plenary session of the Conference of European Statisticians*. New York: Author.

University of Otago – Wellington School of Medical and Health Science (nd). *New Zealand Census-Mortality Study*. Available from:
www.wnmeds.ac.nz/academic/dph/research/HIRP/nzcms/index.html