

IAOS SATELLITE MEETING “MEASURING SMALL AND INDIGENOUS POPULATIONS”

Chungui Qiao
Centre for Social Research and Evaluation
Ministry of Social Development

The IAOS (International Association for Official Statistics) Satellite Meeting “Measuring Small and Indigenous Populations” was held in Wellington on 14–15 April 2005. The meeting was hosted by New Zealand’s national statistics office, Statistics New Zealand, and was held in conjunction with the International Statistical Institute (ISI) 2005 Conference in Sydney, Australia.

Approximately 230 delegates from all over the world attended the meeting, which presented five keynote speakers and another 76 international and national speakers addressing issues and challenges in measuring small and indigenous populations. Tū Williams, Chairperson of the Māori Statistics Forum, led a Māori welcome ceremony for the participants. Brian Pink, New Zealand’s Government Statistician, then opened the conference by identifying the goal of the meeting as being to challenge existing thinking on official statistics – which he declared to be successfully achieved at the conference’s close. Conference abstracts and papers are available from the Statistics New Zealand website: www.stats.govt.nz/about-us/events/satellite-meeting/default.htm.

THE THEME OF THE MEETING

The general theme of the meeting was addressing key issues facing small and indigenous populations, measuring and monitoring their social and economic wellbeing, and exploring practical approaches to progress in these areas. Following the establishment by the United Nations (UN) of a Permanent Forum on Indigenous Issues in 2000, as a subsidiary organ of the Economic and Social Council (ECOSOC), it seemed inevitable that sooner or later an international gathering of statisticians and allied professionals would be convened to explore the dimensions and challenges of measuring small and indigenous populations. Much deliberation preceded this UN initiative, as precise definition of the world’s indigenous populations had been elusive. These crucial issues are directly related to the development of social and economic statistics for the populations of Māori and Pacific peoples, and other ethnic minorities in New Zealand.

While the main focus was on methodological issues, the meeting was also of direct interest to economic and social researchers, as the theme encompassed the main challenges of measuring small domains in economic, social and population statistics.

Other areas of focus included the following issues relating to small and/or indigenous populations:

- statistical methods used to collect information in censuses or other large-scale surveys
- measurement and improvement of data quality
- the derivation of indicators on indigenous populations, communities or subnational areas
- statistical use of administrative data sources
- experience with data integration initiatives
- small area estimation
- other new developments in satisfying the growing demand for disaggregated statistical measures.

Keynote Addresses

The five plenary sessions addressed the overarching theme of the rights and concerns of indigenous peoples, the practical difficulty of collecting information on their social and economic development through traditional approaches, and the obligations and strategies of government departments (especially national statistics offices) in the light of the UN's initiatives on data collection and disaggregation for indigenous populations. The keynote speakers also discussed methodological issues and relevant statistical modelling approaches to meet the information needs of the small and indigenous populations, and presented successful examples in these areas. Papers presented by the keynote speakers included:

- "Political and methodological issues in country experiences of measuring ethnic communities, small and indigenous populations" by Len Cook, National Statistician, Office for National Statistics, United Kingdom
- "Telling stories with numbers: Whose stories? Who is the storyteller? Who listens? Who acts?" by Professor Linda Tuhiwai Smith, Co-Director, Ngā Pae o te Māramatanga (National Institute of Research Excellence for Māori Development and Advancement), University of Auckland, New Zealand
- "When systems collide: On the enumeration of indigenous Australian populations" by Dr John Taylor, Senior Fellow, Centre for Aboriginal Economic Policy Research, Australian National University, Australia
- "Small area estimation of literacy in developing countries with application to Cambodia" by Professor Danny Pfeffermann, Department of Social Statistics, University of Southampton, United Kingdom, and Hebrew University, Israel.
- "Report of the United Nations Workshop on Data Collection and Disaggregation for Indigenous Peoples" by Michael Dodson, member of the United Nations Permanent Forum on Indigenous Issues, Australia.

Concurrent Sessions

The 22 concurrent sessions covered the conference theme using a wider spectrum, varying from poverty and social development issues of ethnic minorities, and measuring and monitoring progress of indigenous populations, to data collection problems and methodologies for these groups. They focused on the following topics:

- measuring indigenous health
- internal migration of Australia's indigenous population
- small area estimation for health planning
- developing a statistical framework for indigenous populations
- population and development in the Pacific islands – rediscovering the relevance
- estimating small area/domain populations – challenges and issues
- Māori development issues
- statistical issues in enumerating small populations in the census
- indigenous businesses and the economy of mana or affection – a Māori case study
- data collection, privacy and confidentiality issues
- official statistics on ethnicity: measurement issues
- collecting data on indigenous populations
- applications of small area techniques
- internal migration – concepts, issues and dynamics
- census of small countries
- small area estimation
- indigenous labour-related issues
- indigenous development issues
- measuring small and indigenous populations – a regional perspective
- indigenous social issues
- issues in enumerating indigenous populations
- Māori health issues.

Main Issues Emerging from the Conference

Many issues with policy implications for the New Zealand public sector emerged from the conference. These range from implementing recommendations of the UN Permanent Forum on Indigenous Issues, to collecting data for small geographical areas, ethnic minorities and other meaningful small domains, to addressing disparity and inequality of social and economic development for indigenous populations. These are of specific relevance to New Zealand in terms of:

- implementing the Local Government Act 2002 by territorial local authorities in policy planning for desirable community outcomes
- monitoring progress in social wellbeing for particular demographical groups, such as Māori and Pacific peoples.

These are briefly summarised in the sections below, which cover general issues, measurements issues, small area estimation and implications for New Zealand.

INDIGENOUS POPULATIONS

The Definition of Indigenous Populations and Their Problems

There are different opinions across regions when it comes to the concept of "indigenous" and the terms "tribal" or "ethnic group". Within the context of the Draft Declaration on the Rights of Indigenous Peoples, it has been decided not to adopt any formal definition of the term indigenous, and self-identification has been stressed by the UN Permanent Forum on Indigenous Issues. The general consensus is that it is crucial to recognise the right to self-identify as part of the right of self-determination, although there are substantial complexities concerning self-identity. In many cases indigenous descent is seen as a prerequisite for indigenous identity, but traditional adoption and self-inclusion within a group should also be considered and respected. Thus no universal definition of indigenous peoples exists for statistical purposes.

From the literature and the lengthy deliberations of the now-superseded UN Working Group on Indigenous Populations, four guiding principles have been generally accepted – indigenous peoples include those who:

- are descendants of the original inhabitants of a country
- have become encapsulated in their lands by a numerically and politically dominant invasive society
- retain cultural difference from that society
- self-identify as indigenous.

For the statistical measurement of indigenous populations, this last criterion – self-identification – is most crucial. However, indigenous populations are of minority status within the data-gathering systems of nation states. If indigenous peoples are to be visible at all in a statistical sense, then administrative mechanisms need to be in place to ascribe and record indigenous status, and indigenous peoples need to respond accordingly and need to be free to do so. It remains an elusive challenge for most countries to form the statistical basis for outlining a precise and truly global description of indigenous populations, as expected by the UN Permanent Forum.

The Workshop on Data Collection and Disaggregation for Indigenous Peoples

The UN Permanent Forum on Indigenous Issues has held three sessions to address issues on priorities for indigenous peoples. The Workshop on Data Collection and Disaggregation for Indigenous Peoples was convened in accordance with ECOSOC decision 2003/300, when data collection was identified as an urgent priority by the

Permanent Forum. The workshop was held on 19–21 January 2004, with 36 experts in attendance from the UN system and other intergovernmental organisations, governments, indigenous organisations, and academia. The workshop discussed a number of case studies, examined challenges and made recommendations concerning data collection and disaggregation concerning indigenous peoples.

Hidden Poverty and Inequality of Ethnic Minority Populations

Progress in social and economic development of the “mainstream” population often leads to the neglect of ethnic minorities or indigenous populations within the same country. The levels of poverty and inequality of certain ethnic minority populations are shocking. Roma, or “Gypsies,” are a good example of this. They are Europe’s largest minority and live in nearly all of the countries of Europe and Central Asia. Europe’s Roma population is estimated at 7–9 million people. Romania is the country with the highest absolute number of Roma, ranging between one and two million. Nearly 80% of the total Roma population lives in European Union candidate countries and new member states. Poverty rates for Roma range between four and ten times that of non-Roma in Bulgaria, Hungary and Romania; nearly 40% of Roma in Romania and Bulgaria live on less than \$2.15 per day.

Roma poverty is multifaceted. Roma are often deprived of the resources necessary for adequate living conditions. Moreover, they lack access to education, labour markets, social and health services, and channels for participation in society. It is not uncommon to find reports of unemployment rates of up to 100% in Roma settlements due to their low skill levels and discrimination in the labour market (they are often the first to be laid off). Education levels for Roma vary markedly within countries, between urban and rural areas, and across different types of Roma communities. In Bulgaria 89% of Roma have primary education or less. Health conditions among Roma are significantly worse than for the rest of the population in most countries. Roma life expectancy is, on average, 10 years less than that of the majority population in Central and Eastern Europe.

MEASURING INDIGENOUS POPULATIONS

The Technical Difficulties of Collecting Information for Indigenous Populations

It is often too hard to collect information on indigenous populations, especially those who live in remote areas with a different culture from the “mainstream” population. For example, some Australian aborigines not only live in remote rural areas, but also have a highly mobile lifestyle. This makes it extremely difficult to reach them for a survey, census or other relevant information collection. Similarly in Guatemala, the indigenous population lives in an extraordinarily isolated situation and speaks up to 13 different

local languages. To conduct a survey of these populations, interviewers must travel on foot for at least six hours before reaching one of the isolated residents. They have to be able to speak the same local language or have access to competent local interpreters in order to measure how the people are living and what their needs are. It is obvious that conventional survey and sampling methodologies would not work for them.

In Australia, prior to the 1971 Census, relatively few resources were applied to the enumeration of indigenous peoples residing in remote areas. For the most part, surveying was restricted to people who by legislation had been "assimilated" as wards of the State on mission and government settlements. Enumeration was largely an adjunct to welfare administration. Since 1971, special census field procedures have been progressively devised, modified and extended by the Australian Bureau of Statistics (ABS), and these now constitute a multifaceted Indigenous Enumeration Strategy designed to account for geographic isolation, high mobility, traditional culture, communication problems due to language and unfamiliarity with form-filling. Key elements of this strategy include interview-based enumeration, use of modified forms, and an extended enumeration period for days (even weeks) around, not on, census night.

In 2001, the direct cost of enumerating remote indigenous populations was around \$2 million (\$26 per head) compared to \$49 million (\$2.60 per head) for the full census. Difficulties arise, however, in attempting to establish the efficacy of these efforts in producing an optimal count, as the normal method of checking for such accuracy, using a post-enumeration survey, is not applied in sparsely settled areas. Despite this, ABS has claimed at times that their methods in remote areas probably yield an overcount. Equally, though, analysts, local indigenous leaders, and service providers have often asserted underenumeration. Separating fact from fiction with regard to these issues is no easy task, but it is true that census counts of usual residence are highly variable over time and space, while even modified enumeration methods struggle at times to accommodate indigenous circumstances, where the concept of usual residence has little meaning.

Counting the Wik in Cape York Peninsula

One study directly tested the validity of a census count for a remote indigenous community, which was an ethnographically based enumeration of Wik Aboriginal people in the community of Aurukun in Far North Queensland, Australia. This enumeration was concurrent with the 1986 ABS Census count, conducted as a conventional count of people found present in dwellings. The enumeration of Wik Aborigines used genealogical methods to develop a comprehensive regional population list of all Wik people whose families were Aurukun residents. This produced a master list of the *de jure* population of the Aurukun region, to be used as a checklist against the census count of the same target population at the same time.

Results showed that the official census fell short of the ethnographically based count by 17%, with most of this gap accounted for by people under 30 years of age. The ABS special procedures had not been introduced in Queensland communities at that time, but an even greater discrepancy was estimated for the 1991 Census after they were adopted. The conclusion was that the enumeration strategy adopted by the ABS for use in remote indigenous communities was structured in such a way as to increase the likelihood of omitting young people, the more mobile and the more socially marginal. The basic problem was, and still is, that this is an attempt to assign individuals to particular dwellings and households, and these are “mainstream” constructs which do not fit the Wik situation, where extended households and high inter-household and inter-community mobility are the norm and the concept of “household” is itself alien to the majority of the population. While there might be cultural differences between indigenous and non-indigenous Australians, the census is not actually aimed at finding out what those differences are. Instead, it is aimed at slotting indigenous Australians into “mainstream” categories. An example of this is provided by the classification of households.

Methodological Issues in Measuring Indigenous Populations

Challenging issues for statistical measurement arise because we often deal with different indigenous and non-indigenous worldviews that can extend all the way from the conceptual and methodological bases for who and what are being measured through to questions of data ownership and purpose. Many of the issues faced by this conference arise from attempts to measure one set of social, cultural and economic systems (indigenous, small scale) using the tools, methods and purposes of another (mainstream, national/global). We also need to address the question of whether simple modifications to the “mainstream” methodologies are sufficient to achieve meaningful measurement of the indigenous population, or whether entirely new approaches are required. We should be mindful that from an indigenous perspective the notion of measurement often carries with it the spectre of state control. The implications of who are measuring what, for whom and to what end, should be crucial points for discussion.

According to the Australian census classification, for example, a household is held to comprise people in a dwelling “living and eating together as a domestic unit”, and the coding of “family types” is derived from a “family reference person around whom a family can be constructed”. Families are implicitly assumed to be mapped onto dwellings, with a structure similar to, or in essence variants of, the “mainstream” nuclear one. These are highly problematic definitions when applied to Aboriginal households, particularly those in remote areas, which are typically fluid in composition, often with a more or less stable core and a variable periphery of transient residents drawn from the same regional population pool. As a consequence, co-residence (even in the limited sense of who sleeps where), family groupings and

domestic economic units are not necessarily coterminous (e.g. people live together, but may not eat together).

The basic economic and social units consist of linked households rather than single ones, and Aboriginal "families" are typically dispersed across a number of households. Commonly such clusters, each comprising five households drawn from a single family group in Aurukun, form the basic units of consumption in remote Aboriginal Australia. The high mobility (especially among teenagers and children) between the constituent parts of such households adds to the complexity of measurement. Compounding the difficulty of establishing household membership is the potential extension to areas outside of the immediate community, essentially to any locality where kin reside. The need for frequent absences to access higher-order services renders the whole notion of a single place of usual residence problematic. One detailed study of mobility in a central Australian desert community showed that up to 35% of the resident population spent significant periods of time away over a given year. In such circumstances, it is difficult to guarantee the accuracy of count without some means of crosschecking census returns against expected usual residents.

SMALL AREA ESTIMATION

Statistical Modelling for Small Area Estimation

Collecting census data for small and indigenous populations, while possible, is much more expensive than what is needed for "mainstream" populations. There are numerous methods for using information from national and/or regional surveys, for which estimates are reliable, to derive small area estimates, with the concept of borrowing strength from other, information-rich, areas. The most recent development is the informative sampling methodology, in which the areas of interest are classified into two categories – sampled and non-sampled areas. Then, appropriate probability theories and application models, plus diagnostic statistics, are used for optimal estimation at small area levels. Bayesian models have also gained popularity as a promising approach that relaxes many of the unrealistic assumptions so that data-based optimal models can be finalised with confidence. The major drawbacks of these "advanced methodologies" are twofold: (1) the modelling generally requires a high level of expertise and experience in this area; (2) the ongoing maintenance of such a modelling process is potentially extremely costly.

An alternative is to resort to a simple cost-effective modelling approach, in which the models can be relatively easily managed by someone with fundamental modelling skills and experience. Loglinear modelling within the generalised linear model framework, for example, belongs to this category and can be considered for such

purposes.¹ The major advantage of this approach is that it enables creation of updated and timely data for the variable of interest at any desirable level of the population, including different small or indigenous populations. Another advantage is that it does not require extremely high levels of expertise for the modelling estimation itself or the regular ongoing maintenance of the estimation. The modelling process itself allows the small area estimates, when summed over any categorical variables, to align with more reliable large area statistics. Naturally, this method may run the risk of misspecification of models, but its overall benefits outweigh those of other modelling options.

In practical applications, the choice of a small area estimation methodology depends on the following factors:

- the drive for gaining small area estimates
- the financial constraints
- the required level of accuracy of statistical estimation
- the availability of statistical modelling expertise for the candidate methods
- the on-going cost in deriving and improving the small area estimates over time.

It is important to note that once a particular method has been chosen, it is best to stick to the same method in order for the small area estimates to form consistent time series data. Frequent changes of estimation methods are not considered good practice. A practitioner may aim for more realistic estimates with intensive resource investments over time, or resort to some simpler methods with lower running and maintaining maintenance costs. The decision is sometimes a difficult one to make.

Political and Confidentiality Issues

Small populations can involve all people in a small state, or subgroups of a large entity, defined by some demographic or social attribute such as gender or ethnicity, or even small groups of businesses and other economic entities. Government agencies seek to identify common themes and developments, and explain how policy and politics interact. Availability of sound statistical measures is not determined solely by funding or even political acceptance, but also depends on relevant methodologies and addressing issues of privacy and confidentiality. It can require official statisticians to develop an understanding of cultural issues that is often difficult to achieve. In some areas, the production of official statistics may create tension between the independence of the statistician and the authority of the state.

1 The author, Chungui Qiao, presented a paper on "Small area estimation of unemployment using loglinear models in SAS" at the meeting, and the full paper has been accepted for publication in *Labour: Review of Labour Economics and Industrial Relations* (2005) Vol 19, Issue 4.

Financial Constraints

Funding is always a constraining factor for collecting data on small and/or indigenous populations, apart from the difficulty of physically reaching these people. For obvious reasons, it is not feasible and affordable to frequently measure these populations directly, as in the population census. The budget is limited even for small area estimation in the United Kingdom, where it has been found extremely costly to model the estimates statistically for small populations and to adjust the models on a continual basis. The cost of direct measurement of these populations would be even more expensive. Thus, it appears sensible to strike a balance between gaining sufficient statistical estimation accuracy and keeping the budget within financial capability.

IMPLICATIONS OF INDIGENOUS ISSUES IN NEW ZEALAND

Māori and Pacific Peoples

Public policy in New Zealand has a specific focus on addressing disparity and inequality in social development between Māori and Pacific peoples and the "mainstream" European population. Statistics New Zealand has addressed this issue by making sure that statistical needs regarding Māori and Pacific peoples are more clearly identified and progressively met in the design of surveys, census and other data-collecting activities. The four aspects of this work are:

- improving the relevance of official statistics to Māori, and Pacific peoples
- raising awareness about, and extending the use of official statistics in Māori and Pacific communities
- building statistical capability in Māori and Pacific organisations
- building internal capability to respond to Māori and Pacific needs.

It will probably continue to be a challenge for the government to raise public awareness of these issues and to gain common support in this area.

Community Development under the Local Government Act 2002

Through the Local Government Act 2002, territorial local authorities are required to develop their Long-Term Council Community Plans with desirable community outcomes to be defined and continually monitored. Here the communities could refer to any subset of the New Zealand population, small populations or indigenous populations. Developing these plans and monitoring systems not only requires relevant and timely data for these populations, but also an understanding of how such data could be best utilised to inform policy decisions at local levels. The data gaps at the territorial local authority or other small area level (including community level) present a major problem for the implementation of this Act among the local stakeholders.

Administrative data from central government agencies can serve as a valuable source of information for local government and communities. Although these data have not been made easily accessible, a few government agencies have taken initiatives in this direction. Examples of these initiatives include recent progress in the Quality of Life Report for the Big Cities project, disaggregation of some of the indicators used in the Ministry of Social Development's *The Social Report* to subnational level, and the Linked Indicators Project to provide core indicators linking national and local information on social development. Data integration to combine data from different sources and statistical modelling to generate new data for targeted audience have both shown promise in meeting local information needs. More remains to be done to provide the relevant data on the right population of interest on a continual basis.