



Australian Government

Australian Institute of Health and Welfare

Australian Institute of Family Studies

Closing the gap clearinghouse

Data about and for Aboriginal and Torres Strait Islander Australians

Issues paper no. 10 prepared for the Closing the Gap Clearinghouse

Nicholas Biddle

July 2014

Contents

Summary	2
What we know	3
What we don't know	3
1 Introduction and overview	4
2 Users and uses of data about Indigenous Australians	6
2.1 Government and other policy makers	6
2.2 Academics and researchers	10
2.3 Indigenous community and other organisations	12
2.4 The wider Australian community	13
3 The current statistical environment for data about Indigenous Australians	14
3.1 Cross-sectional surveys	14
3.2 Longitudinal collections	17
3.3 Qualitative data	20
3.4 Administrative data	21
3.5 Community-controlled collections	24
4 Gaps between data needs and data availability	26
5 Concluding comments and how Indigenous data could be improved	29
References	33



Summary

This paper provides a summary of who uses data about and for the Aboriginal and Torres Strait Islander population of Australia and how these data are used:

- Indigenous community and other organisations use data for reflection and advocacy.
- Government and other policy makers use data for allocating resources, identifying needs, monitoring, evaluation and planning.
- The wider Australian community uses data for accountability (of politicians and policies) and to provide more information about the circumstances of non-Indigenous people.
- Academics and researchers use data for understanding and explaining.

This paper discusses the availability, strengths and weaknesses of 5 types of data: cross-sectional surveys, longitudinal data, qualitative data, administrative data and community-controlled collections.

The extent to which the data sets and types of data meet the uses of the data varies considerably. While the needs of governments and researchers are generally well provided for, the needs of Indigenous community and other organisations who work with and for Indigenous Australians are not as well met due to their reliance on community-controlled collections that are relatively poorly resourced. Other gaps and issues include:

- complexities around identification, where the same individual could be recorded as having a different Indigenous status across collections or through time
- the lack of a data set with longitudinal information on Indigenous Australians across the life course (that is, from childhood through to adulthood)
- the use of a range of data sets for a single purpose, meaning that scope and definitions may not be consistent
- the reliance on the Australian Census, a data set that is the only source of data for some of the uses discussed above (in particular allocation and planning) and is collected only every 5 years
- difficulties in accessing administrative and other data by those outside government, although large tables of appendices from some of those data collections are published
- with some exceptions, the shortage of data at the right level of geography—by socially meaningful groupings (for example women and men or across the life course), for the right population and for the right concepts—to be useful for Indigenous community and other organisations
- the lack of trial and program data that will enable the evaluation of the effectiveness of programs and policies.

With these limitations in mind, this paper discusses 4 suggestions for improving data about and for Indigenous Australians:

- considering the inclusion of Indigenous specific information in the Census
- expanding the availability of longitudinal data, ideally through the linking of existing cross-sectional data sets
- enhancing evaluations (including of service availability and effectiveness) with funding built into the design of programs at the initiation phase and making unit-record data publicly available
- increasing support for Indigenous quantitative researchers through targeted scholarships and a national training program for Indigenous data collectors, data managers and analysts.



What we know

- The Indigenous population have a greater likelihood of being involved in quantitative and qualitative research than most other Australians.
- We have reasonable information about where Indigenous people live and the distribution of their socioeconomic outcomes, albeit with some error and uncertainty.
- We know that there is a large gap between Indigenous and non-Indigenous Australians in health and socioeconomic status, and that the gap holds across Australia.
- A consistent finding from data about Indigenous Australians is that there are different levels and determinants of wellbeing depending on the measure used. For example, on some measures of wellbeing, those in remote areas do relatively poorly, in others they do reasonably well.

What works

- From both a rights-based and pragmatic perspective, the inclusion of Indigenous people and Indigenous standpoints in the design and analysis of Indigenous-specific data collections and surveys enhances their relevance and acceptance within the Indigenous community.
- The consistent use of Indigenous identification questions in administrative and other data allows for comparability across data collections.
- Combining qualitative and quantitative data gives a more nuanced view of Indigenous outcomes and aspirations.

What doesn't work

- Government program and policy evaluations that are designed after the policy has been implemented are likely to be more expensive and usually do not give useful information about the effectiveness of the policy.
- If we assume that a particular policy or program will be effective, then we are unlikely to learn much from its evaluation.
- Assuming that Indigenous people will respond to surveys or stay in longitudinal surveys at the same rate or for the same reasons as the non-Indigenous population can lead to high rates of non-response and sample attrition.

What we don't know

- We rarely consider or discuss how to balance the competing priorities of the different uses and users of data on Indigenous people.
- Due to the lack of longitudinal information, we know little about the specific causes or consequences of changes in Indigenous outcomes.
- There is a consistent lack of information and data on the specific policies and programs that will lead to sustained improvement in Indigenous wellbeing or what some of the unintended consequences of particular programs might be.



1 Introduction and overview

Data, or information collected and organised for analysis and interpretation, can shape debate and guide policy decisions. Data can be used to identify areas and populations with the greatest need and direct resources accordingly. Data allow us to monitor change and show where outcomes are improving, stagnating or worsening. Most importantly, data can be used to evaluate policies and programs for their efficacy and cost effectiveness, thereby improving individual and collective wellbeing.

However, data collections that are poorly designed or poorly targeted can distort the policy process and lead to inefficient outcomes. A data set that is not collected in consultation with the population that it focuses on can privilege outcomes of minimal relevance while marginalising outcomes that are of greater importance to that group.

The methods and processes for collecting data are of particular importance for Aboriginal and Torres Strait Islander people. Collectively, Indigenous people of Australia were estimated in June of 2011 to have a population of just under 670,000, or 3.0% of the total Australian population (ABS 2013a). Despite Indigenous Australians making up a small share of the total Australian population, the need to collect and analyse data to understand the demographic, socioeconomic and health dynamics of the population far exceeds their population size.

Fiscally, the allocation of funds from the Commonwealth to state and territory governments is influenced in part by how many Indigenous Australians are estimated to be living in that jurisdiction (CGC 2010). Related to this, at both the national and state and territory levels, the Council of Australian Governments (COAG) has devoted considerable resources to a set of policies with the aim of 'Closing the Gap' (Australian Government 2010). These policies include a target to eliminate the disparity in life expectancy between Indigenous and non-Indigenous Australians within a generation. However, the 'Closing the Gap' policy framework also includes targets related to child health, education and employment, with ongoing discussions related to new targets like school attendance, disability or criminal justice. Monitoring of progress towards these targets requires detailed data collection and careful analysis.

Data about Indigenous Australians are also crucial at the local level. For governments providing services that are specific to or tailored towards Indigenous people, as well as those providing mainstream programs, it is crucial to know where Indigenous Australians live as well as how the local Indigenous population is changing and expected to change. Furthermore, governments and service providers also need to know the characteristics of that local population, including the age composition and socioeconomic outcomes. Indigenous Australians make up a small percentage of the overall Australian population (3%). Most Indigenous Australians live in major cities (35%) and regional areas (44%), while 22% live in remote and very remote areas. This means that with the exception of some concentrations of the Indigenous population in relatively remote areas, in much of Australia, the Indigenous population forms a small minority. Even so, Indigenous Australians also make up a disproportionate share of certain policy-relevant age groups, including pre-schoolers, school-age children and young adults, those entering the workforce, and those entering their peak childbearing years.

One of the uses, if not the most important use, of data about the Indigenous population is giving Indigenous people the power to tell their own story and advocate for change. Data have the potential to empower communities to shape their own future. Once again though, this will be most effective if the data are collected either in collaboration with the relevant Indigenous population or by the Indigenous population with sufficient resources.

There are many data collections that focus on the Indigenous population. These include the National Aboriginal and Torres Strait Islander Social Survey (NATSISS), the Australian Aboriginal and Torres Strait Islander Health Survey (AATSIHS, formerly the National Aboriginal and Torres Strait Islander Health Survey (NATSIHS)) and the Longitudinal Study of Indigenous Children (LSIC). Other collections, like the Census of Population and Housing and the Longitudinal Surveys of Australian Youth (LSAY) have specific strategies for collecting data about



Indigenous people, or in the case of the LSAY, have an over-representation of Indigenous people in the data set. There is also a small but emerging set of data about Indigenous people collected by Indigenous community and other organisations for the evaluation of particular policies or programs. Finally, most government administrative data sets include a voluntary Indigenous identifier, and data about the Indigenous population can be extracted, for example Medicare and hospital data. In total, there are more than 50 national data collections that include data about Indigenous Australians.

Leaving aside administrative data (where most of the reporting burden falls on government), there is a large respondent burden on the Indigenous population in terms of participation in government surveys or academic research. There is also an additional burden for those that have a role in representing and advocating on behalf of particular communities in terms of data collection. While this advocacy is important (and discussed later in the paper), it makes it particularly important that data is collected in the most efficient way possible with maximum use made of each data collection.

Despite this array of data about Indigenous Australians, there are many gaps in our understanding of Indigenous demography, health, socioeconomic status and wellbeing. As will be discussed in this paper: data are often not available in a comparable way for the non-Indigenous population; data are often out of date and not reflective of current circumstances; the quality of the data is often suspect; there are some gaps in information on the change in the outcomes of individual Indigenous Australians through time; data are often not available at the level that is needed by Indigenous communities; Indigenous-specific measures of wellbeing are often missing; and there is little to no information on the effect of particular policies or programs on Indigenous Australians.

A final issue relates to the issue of identification. For very good reasons, Indigenous status in Australia is primarily based on self-identification, with most data collections using a single question. Furthermore, this question is often answered by someone else on behalf of the individual. Significant improvement has been made in how consistent this question is across data collections. While the question may be constant, the context within which it is asked is not. This means that an individual may be recorded as having a different Indigenous status across collections or through time, making comparisons quite complex.

It needs to be considered, therefore, whether there are ways to collect data about Indigenous people that better meet the needs of government, researchers and most importantly the Indigenous community. Using a review of the available literature both within Australia and abroad, this paper will attempt to identify best practice in terms of data collection on small self-identified populations, as well as the stated needs and desires of government, researchers and Indigenous communities for data about Indigenous Australians.

Structure of the report

The report is structured as follows:

- Section 2 discusses the potential users and uses of data about Indigenous Australians. Potential users will be broken into 4 broad groups: government and policy makers; researchers and the academic community; Indigenous community and other organisations; and the wider Australian and international community.
- Section 3 gives an overview of the current statistical environment for data about the Indigenous population of Australia and focuses on the availability and limitations of cross-sectional data sets, longitudinal data sets, administrative data and community-based collections.
- Section 4 discusses the extent to which existing data are adequate for the desired uses.
- Section 5 summarises the issues and some suggestions for improvements to data about Indigenous Australians based on best practice from Australia and abroad.



2 Users and uses of data about Indigenous Australians

Data that are collected with information on Indigenous Australians are used by a variety of stakeholders and for a range of purposes. There is considerable overlap with, for example, governments and Indigenous community organisations both having an interest in knowing whether the outcomes of the Indigenous population in a particular community are improving or worsening; but there is often a tension between different users. Extending the example, the outcomes that an Indigenous community organisation might be interested in monitoring might be very different to those that the government is interested in. With a finite budget and, perhaps more importantly, a desire to minimise respondent and provider burden, there will therefore be trade-offs that need to be made.

The aim of this section is to discuss the current and potential uses of Indigenous data. These uses are allocated to their primary users (while recognising significant overlap): government and policy makers; researchers and the academic community; Indigenous community and other organisations; and the wider Australian and international community. Throughout the section, I discuss explicitly some of the main tensions that arise between these different uses and users.

2.1 Government and other policy makers

All levels of government in Australia use data about Indigenous Australians, albeit for slightly different purposes. Furthermore, the Commonwealth Government, and to a lesser extent, the state and territory governments, are the primary data collectors. They collect data directly through their Census and survey programs, by funding other organisations to collect the data for them, or through the creation of large administrative databases. The main ways in which they use this data are outlined below.

Allocation

The Government Financial Estimates compiled by the Australian Bureau of Statistics (ABS 2013b) include an estimate of total revenue across all government sectors of \$553.5 billion for the 2013–14 financial year. While most of these funds are spent on items not specific to the Indigenous population, the Steering Committee for the Review of Government Service Provision (SCRGSP 2012) estimated that, for the 2010–11 financial year, direct Indigenous expenditure equalled \$25.4 billion, or 5.6% of total direct government expenditure for that year. Accounting for this expenditure requires accurate estimates of the size of the Indigenous population as well as an identification of the Indigenous status of those who use particular services.

In addition, the Commonwealth Grants Commission (CGC) recommends how the revenues raised from the goods and services tax should be distributed to the states and territories to achieve horizontal fiscal equalisation. In the 2010 Review Report, the CGC (2010:7) states that ‘a major difference across states arises because of the uneven distribution of Indigenous Australians (particularly in more remote locations), and the higher expenditure states incur in providing services to these residents.’ The allocation of goods and services tax revenue is therefore influenced by the proportion of Indigenous Australians that live in each state and territory, thereby necessitating an accurate estimate of the population in each jurisdiction.

Need

One of the reasons Indigenous status is used to allocate resources in some instances is that the cost of providing services to the Indigenous population is often higher than the cost for the non-Indigenous population. There is also substantial diversity within the Indigenous population: many Indigenous Australians have very similar needs to the rest of the Australian population, but others have more expensive or more complex needs. It is necessary, therefore, to take into account the level of absolute and relative need in the Indigenous population.



Several authors have created summary measures of socioeconomic outcomes for the Indigenous population in attempts to identify this need. In the early 1990s, Tesfaghiorghis (1991) used 3 variables derived from the 1986 Census—representing education, employment and income—to construct an index of socioeconomic advantage at the Aboriginal and Torres Strait Islander Commission Region level for the Indigenous population.

Altman & Liu (1994) used a similar list of variables to examine socioeconomic status for a reduced number of regions for the 1991 Census, making some comparisons with results from the 1986 Census. Gray & Auld (2000) constructed an index of relative disadvantage using 4 variables that represent family income, housing, educational attainment, and employment. The authors found a reasonable level of stability between 1991 and 1996 in terms of how regions ranked, with Alice Springs and Cairns being notable exceptions. Biddle (2009a) found a similar result for the 2001 and 2006 Census, making comparisons all the way back to the analysis of Tesfaghiorghis (1991).

There have been a number of innovations in this index-based analysis of the Census. Gray and Auld (2000) augmented their Census-based analysis with administrative data to attempt to control for the Community Development Employment Program scheme. In the same year, the CGC funded the ABS for a major study to construct indexes of Indigenous socioeconomic disadvantage (ABS 2000). One major difference between ABS (2000) and the previous studies was the combining of data from the 1996 Census and both the 1992 National Aboriginal and Torres Strait Islander Survey (NATSIS) and perinatal statistics. Another difference was that 9 separate indexes were created representing different aspects of socioeconomic advantage or disadvantage.

Biddle (2009b) created a pooled index that enabled intra-area comparisons between Indigenous and non-Indigenous outcomes; the results were updated for the most recent Census in Biddle (2013). This most recent analysis also introduced a Suite of Indigenous Outcome Indices to compare area level ranking across 5 domains: Income and wealth; Employment; Education; Child outcomes; and an Index of language, social and health maintenance. Finally, Yap & Biddle (2010) created a separate set of indices to compare and contrast the different socioeconomic distribution of Indigenous men and women.

While the methods and focus of this identification of need have varied, the findings have tended to remain consistent. First, there is considerable diversity within the Indigenous population in terms of socioeconomic outcomes, both by geography and by gender. Second, geographic patterns are reasonably consistent through time. Third, there is a strong correlation between dimensions of socioeconomic status but a weaker correlation with health or other broader notions of wellbeing. Fourth, child outcomes correlate weakly with adult outcomes at the geographic level. Finally, the analysis has shown that, regardless of the measures used, Indigenous Australians in an area tend to have worse socioeconomic outcomes than non-Indigenous Australians.

Monitoring

Socioeconomic indices have tended to focus on comparisons within the Indigenous population either by geography, gender or type of socioeconomic measure. At least since the formalisation of the Closing the Gap targets—and even earlier under previous Australian Government policies of Practical Reconciliation and the Aboriginal Employment Development Policy—the focus of government targets has tended to be on the position of Indigenous Australians relative to the rest of the Australian population. A large amount of data and analysis is therefore devoted to monitoring whether progress is being made in meeting targets at the national and sub-national level. The biennial *Aboriginal and Torres Strait Islander Health Performance Framework* is a notable example in the health sector (Box 1).

The most extensive analysis undertaken on this monitoring is the *Overcoming Indigenous Disadvantage* (OID) report produced every 2 to 3 years by the Steering Committee for the Review of Government Service Provision, with support from a Secretariat located in the Productivity Commission. The most recent report (SCRGSP 2011) used a combination of survey and administrative data to look at 45 quantitative indicators. The next report is likely to rely more heavily on data from the 2011 and previous Censuses, as well as the 2012–13 AATSIHS. It is still, however, likely to be structured around a similar framework of outcome indicators. While the report is extensive,



in the foreword to the report, the Chair of the committee noted that ‘Despite ongoing improvements in data collections, for one third of the indicators in this edition, adequate data were not available to measure changes over time’ (SCRGSP 2011:iii).

Supplementing the OID report and focusing more closely on government targets, for the 5 years leading up to 2013, the Australian Prime Minister has reported to Parliament on progress in Closing the Gap. The latest report was released on 11 February 2014, and it has a much smaller focus on data and a greater focus on priorities and aspirations than previous reports did. The report is based on the Closing the Gap targets that were identified under the National Indigenous Reform Agreement in 2009 and committed to by all Australian governments at the time. One of the roles of the COAG Reform Council was to report on the performance of governments under national agreements. Like the OID report, data availability limits the amount that can be said each year with regards to most of the targets, and the focus therefore tends to be on describing specific programs and budget outlays.

In addition, government agencies also monitor targets related to the Indigenous population. This makes sense because the administrative data required to assess many of the targets are often held by government. Even if the agencies are separate from the executive government and individuals involved operate to the highest ethical standards, government reviews and reports could be perceived as lacking a certain degree of independence and therefore focus on outcomes or measures that paint the government of the day in a relatively favourable light. Public confidence in the monitoring of government targets therefore requires external validation and critique, which in turn requires wider access to administrative and other databases.

Box 1: Health Performance Framework

The Aboriginal and Torres Strait Islander Health Performance Framework (HPF) was developed under the auspice of the Australian Health Ministers’ Advisory Council to inform policy analysis and program development in Indigenous health.

The HPF incorporates data from over 50 sources (including Hospital Morbidity, Mortality, the Census, the National Aboriginal and Torres Strait Islander Health Survey, and the National Aboriginal and Torres Strait Islander Social Survey) in more than 60 indicators included in the framework. The HPF framework covers 3 tiers of health performance—health status and outcomes, determinants of health, and health systems performance. Reporting against the HPF is designed to promote accountability, inform policy and research, and foster informed debate.

The Australian Institute of Health and Welfare (AIHW) receives funding from the Department of Health and the Department of Prime Minister and Cabinet to produce a detailed analyses HPF report every 2 years (AHMAC 2012).

Evaluation

While it is important to track and critique progress in meeting targets related to Indigenous outcomes, such national or even jurisdictional analysis tells us very little about the effectiveness of specific policies and programs in achieving their stated aims. Even within the OID report mentioned previously, there has been an undue emphasis on cataloguing disadvantage as opposed to specific strategies to overcome it. This is not necessarily the fault of those who put together the OID and other reports, as there is a real need for data about Indigenous Australians that can be used to test how effective particular programs and interventions are in achieving their aims.

The Closing the Gap Clearinghouse (CtGC 2013:1) was set up to ‘bring together research and evaluation evidence on overcoming disadvantage for Indigenous Australians’ with a focus on what we know, what works, what doesn’t work, and what we don’t know. It would be fair to say, however, that many of the publications produced for the CtGC conclude that there is a great deal that falls into the ‘what we don’t know’ category. For example, in their review of early learning programs for Indigenous children, Harrison et al. (2012:2) state that ‘There have been no rigorous trials or evaluations of early childhood programs in Australia, particularly programs for Indigenous and at-risk children’. Helme and Lamb (2011:2) make a similar conclusion in their review of evaluations of programs related to school completion and state that ‘There is insufficient evidence demonstrating the effectiveness of many programs that are designed to improve the outcomes for Indigenous young people.’



Evaluations of the way in which programs are implemented are more common. As there is likely to be some variation at the local level in how effectively government programs are implemented, evaluations of these types are very important. For example, the Sentinel Sites Evaluation found wide variation at the local level in how effectively the Indigenous Chronic Disease Package (ICDP) was implemented. Organisations with strong leadership and capacity were more successful in implementing the ICDP measures than organisations that were building their capacity (Bailie et al. 2013). However, there are far fewer evaluations of the effect of government policies or programs.

In October 2012, the Productivity Commission convened a roundtable on 'Better Indigenous policies: the role of evaluation'. This roundtable had around 40 participants from across government, academia and the community sector with the proceedings documented across 13 chapters in PC (2013). In the introduction to the roundtable, the then Chair of the Commission stated that 'Social policy is notoriously difficult to design and evaluate, and Indigenous policy is as difficult as any' (Banks 2013:18). In addition to the often politicised nature of Indigenous policy, a further difficulty is that it is necessary to have data on those who did participate in a particular program (ideally before and after their participation) as well as data from an otherwise identical group of people who did not participate. The latter is obtainable if participation in the program is randomly assigned either intentionally or through natural processes (NHMRC 2000). This is, very rarely the case.

More common is an econometric-regression style of analysis where the outcome of interest is assumed to be a function of participation in a particular program and of other observable characteristics. When looking at the differences between those who did and did not participate in the program, those other observed characteristics are held constant. A similar technique is propensity-score matching where individuals (or groups of individuals) are matched in pairs based on observable characteristics, with the only difference between the pairs being whether or not they participated in the program or received the treatment. While these techniques are often all that is available, the problem with econometric-regression and propensity-score matching techniques is that it is not possible to control for unobserved characteristics, including most problematically, those that are associated with both participation in the program and the outcome of interest. For a useful and accessible discussion on such issues, see Leigh (2010).

The above are data issues that ideally need to be carefully considered in the design phase of policies and programs for Indigenous Australians, not after the policy or program has been implemented. There also needs to be a program logic that is realistic and well-articulated. If it is important to know whether a particular program affects a particular outcome, then it is necessary to adequately control for selection into the program. There are, however, Indigenous-specific considerations with regards to evaluations that, while adding a degree of complexity, have the potential to improve our understanding of what works.

The Co-Chair of the National Congress of Australia's First Peoples stated in his chapter of the roundtable proceedings that 'the United Nations Declaration on the Rights of Indigenous Peoples is the foundation, if not the framework, on which goals, targets and evaluation should be built' (Malezer 2013:69). In particular, evaluations need to be built around the rights of Indigenous people to determine their needs and the means to achieve them. In this way, programs (and the corresponding evaluations) are much more likely to be supported by the Indigenous people that they are designed for; the programs are also more likely to be focused on the outcomes and goals of Indigenous people.

This legitimate need for 'prior informed consent on any programs that may affect [Indigenous Australians]' (Malezer 2013:74) does, however, create a tension. If consent is truly to be a binding constraint, then even if that consent is obtained at the community level, selection into the program will not be random. There are likely to be unobserved factors (such as community capacity) that affect that consent and the outcomes being achieved. Alternative evaluation methods and, therefore, alternative data collection strategies to standard randomised controlled trials need to be considered. One alternative is to randomise the timing of admission into a program (pipeline comparisons) from within consenting communities (Ravallion 2008).



Another alternative form of evaluation that has the potential to be of less ethical concern to Indigenous people and government service providers is randomised promotion. Covered in Gertler et al. (2010), such techniques rely on different individuals or communities being actively made aware of or notified of the potential benefits of a particular program at different levels of intensity. If these differences are large enough and randomly distributed, and people are sensitive enough to promotion and potentially to incentives, then this variation in promotion can be used as an instrumental variable to control for selection into the program.

Planning

Improving Indigenous outcomes requires knowledge not only of where the current Indigenous population lives and what its characteristics are, but also what the likely structure of the future Indigenous population is. For this, we are reliant on population projections or forecasts.

Accurate projections of Indigenous populations can be difficult to obtain (Biddle & Wilson 2013), and the difficulties are heightened at the local or community level at which data are often required. First, there is uncertainty in accurately establishing the base population. Second, the administrative data used to derive estimates of births and deaths are collected with considerable error, a still unknown proportion of Indigenous births and deaths may not be recorded as such. Thirdly, and most importantly, there are changes across time and across collections in the way in which a person identifies as being Indigenous. Not all those who have Indigenous ancestry identify themselves as such, with some people not being aware of their Indigenous ancestry and others being reluctant to identify as such to data collectors. According to Biddle & Wilson (2013:105), 'This change in identification may be in response to wider social norms and pressures, changes in the financial incentive or disincentive to identify, and variation in collection methodology.' In addition, in administrative collections, there is sometimes a tendency for staff not to ask the Indigenous identification question and either assume that the categories of 'non-Indigenous' status or 'not stated' apply.

To demonstrate the extent to which identification may change, it is worth considering the results from the most recent intercensal period. In 2006, the Indigenous population was estimated to be around 517,000. The ABS (2009) projected that by 2011 the population would reach around 575,000. This represented a more rapid growth than that projected for the non-Indigenous population. However, population estimates based on the 2011 Census resulted in a substantially higher population still—around 670,000 Indigenous Australians. Only around 37.9% of the growth over the period was captured in the projection model. Ultimately, obtaining more accurate projections would necessitate incorporating a more rigorous understanding of identification change into projection models. This would require analysing a more sophisticated set of data that has information on Indigenous status at many different points in time.

2.2 Academics and researchers

By far the biggest users and producers of data about Indigenous Australians are government agencies. Academics and researchers also have an interest in the issues of allocation, need, monitoring, evaluating and planning uses of data discussed above. Often, this is through research commissioned by government, but it can also involve independent research that is not directly related to government-funded contracts or centres. There are, however, 2 additional uses of data about Indigenous Australians that academics and researchers focus on—understanding and explaining.

Increasingly, Indigenous researchers are doing this analysis. The concluding section of this paper recommends that increased support be given to quantitative researchers who identify as being Indigenous (though not at the exclusion of qualitative researchers) to ensure the questions being asked and techniques used to elicit answers to the questions are of relevance to the Indigenous population.



Understanding

One of the roles of universities and research centres the world over is to improve the understanding of the society in which they are located, the people that they are serving, and the relationship they have with the wider international community. This is done through both the teaching of undergraduate and postgraduate students as well as through research and scholarly output. Within Australia, therefore, there is a need for data on the Indigenous population to understand the patterns and diversity within the Indigenous population.

In terms of teaching, a review of the websites for 39 major universities in Australia found that 27 have some form of Indigenous studies as a major (usually in a Bachelor of Arts program) or as a stand-alone degree, with an additional 3 universities offering Indigenous studies as a minor only. This leaves only 9 universities that have no option of an Indigenous-specific degree or as a major or minor. The content of the courses vary considerably across the universities and across disciplines, but all are likely to use to a certain degree of data on Indigenous Australians as part of the teaching program, either in raw form or through published research.

Indigenous studies degrees are a relatively recent phenomenon. However, there is a long history of quantitative data-heavy research including Charles Dunford Rowley's series, *Aboriginal policy and practice*, in the early 1970s (Rowley 1970, 1971a, 1971b), Altman and Nieuwenhuysen's (1979) book *The economic status of Australian Aborigines*, and the *Report of the Committee of Review of Aboriginal Employment and Training Programs* (CRAETP 1985). More recently, papers or sets of papers have been built around specific data collections, including the Western Australian Aboriginal Child Health Survey summarised in, for example, Blair et al. (2005), the 2008 NATSISS summarised in Hunter & Biddle (2012), and the most recent Census through the Indigenous Population Project: 2011 Census Papers series (see <<http://caepn.anu.edu.au/publications/censuspapers.php>>).

These (and many other) books, papers and series have inevitably taken a somewhat more critical approach to past and current policy settings than is the case with the reports mentioned earlier (see 'Monitoring' on page 8 of this paper). However, they have been less extensive in the range of data that they have used. This is mainly due to access issues: they have tended to focus on publicly available data, which are limited either by access and reporting (in the case of administrative data sets) or confidentiality and timing (in terms of survey or Census data).

Explaining

One of the focuses of academic research related to Indigenous data is explaining relationships between variables and domains. More so than government statistical reports, academic analysis has focused on the determinants and correlates of relevant outcome variables, with the aim of explaining some of the reasons for variation within the Indigenous population or between Indigenous and non-Indigenous Australians. For example, researchers have covered in some detail the determinants of Indigenous employment and labour market outcomes and the determinants of Indigenous health.

Labour economists are particularly interested in the extent to which observable human capital characteristics explain differences between 2 population groups in labour-market-related outcomes. In a set of techniques that is often referred to as Oaxaca-Blinder decompositions (following Blinder 1973 and Oaxaca 1973), the difference between 2 groups in a particular outcome is decomposed into that which is explained by observable characteristics (such as education and work experience) and the residual unexplained component. This residual component can then be taken as prima facie evidence for the ongoing existence of labour market discrimination. Examples of this or similar analysis include Ross (1993), Hunter (2006), Chapter 6 in Biddle and Yap (2010b) and Kalb et al. (2012). All these analyses, however, have suffered from missing data related to either the quantity or quality of a person's education or their labour market experience. Not being able to control for these differences can lead to an overemphasis on discrimination as opposed to human capital development in explaining Indigenous labour market outcomes.

Alongside employment and income, the determinants of Indigenous health have been a particular focus of analysis of Indigenous data. Similar to the Oaxaca-Blinder decompositions mentioned above, the approach taken has usually been to take a broad measure of health (for example self-assessed health or chronic disease)



and look at the difference between Indigenous and non-Indigenous Australians before and after controlling for observable characteristics. Booth and Carroll (2008) is one example of such an analysis, and they find that almost half of the gap in health outcomes can be explained by observable economic characteristics.

Many other subjective and objective outcomes have been compared in analyses similar to those above. They all suffer from one crucial data issue in that the outcome of interest is observed at the same time as the control variables or factors that are assumed to explain it. This makes it difficult to make policy conclusions based on the analysis because causal pathways cannot be robustly established. Data that measure outcomes at more than one point in time and that compare changes in the control variables with changes in the outcome would be more useful. With this type of data, time invariant unobserved characteristics could be held constant.

2.3 Indigenous community and other organisations

Most users of data about Indigenous Australians are not Indigenous. This includes people from within government and academia. However, the title of this paper makes it clear that an equal focus should be on data for Indigenous Australians, not just about Indigenous Australians. A genuine commitment to providing data in a way that meets the needs of Indigenous community and other organisations might, however, involve considerable trade-offs with some of the uses mentioned above. These trade-offs cannot be ignored but must instead be grappled with.

Reflection

A persuasive argument for 'The power of data in Aboriginal hands' was made by Yu (2011:7) in his role as the Chief Executive Officer of Nyamba Buru Yawuru Ltd, Broome (he was also a member of the Australian Statistics Advisory Council of the Australian Bureau of Statistics at the time). He argued that 'Whilst the power over data collection and analysis remains in the hands of government the narrative about Indigenous people's place in the Australian nation will continue to be one of deficit, disadvantage and dysfunction.'

Jordan et al. (2010) argued that the outcomes of interest to Indigenous community and other organisations can be very different to the outcomes of interest to government. As Yu (2011:4) indicated, the major issue is that government policy and the data that underpin it emphasise 'Indigenous individual achievement in education, employment and home ownership, informed by an ideological assumption that this is what promotes social and family functioning and good physical and mental health'. While important for many Indigenous people and in certain contexts, such data do not 'measure seriously the fundamental imperatives of Indigenous life' (Yu 2011:4). While there is no homogeneous Indigenous viewpoint, imperatives that take on a greater focus might include a group or family orientation towards economics, aspirations and support, as distinct from individualistic orientations. Rowse (2012) has made the point that government's focus on equality of outcomes does not allow Indigenous people to decide for themselves how they will engage with wider society.

In addition to the content of the data collected, a commitment to collecting and providing data in a way that meets the needs of the Indigenous population would recognise that the geographic and social grouping of the Indigenous population by Indigenous people themselves does not necessarily align with the jurisdictional, statistical and administrative groupings that are created by government and used by academics. The extensive and increasing holding of a significant proportion of the Australian continent under some form of Indigenous land tenure highlights a 'growing discrepancy between the best-intentioned of statistical output frameworks and the actual needs of Indigenous land-holding groups for an ethnographically informed demography suited to their needs for managing the Indigenous estate and its associated constituencies' (Taylor 2013:127). Different data, or at the very least different analysis of data, are therefore required when Indigenous people attempt to examine how a particular community or people is progressing.



Advocacy

Indigenous people are increasingly demanding and using data to understand what is happening within their own community. However, there has been an equally important use of data to question whether particular communities or people are receiving the level of government support that they are entitled to. This advocacy role requires quite different types of data, with a focus on the level and quality of services and not just on outcomes.

An example of this use of data is the work by Taylor & Stanley (2005) in the Thamarrurr region of the Northern Territory. Using the *Report of the Royal Commission on Aboriginal Peoples* (RCAP 1996) in Canada as a basis, the authors used an extensive array of outcome and expenditure data to show that there was 'a structural imbalance in funding at Thamarrurr with lower than average expenditure on positive aspects of public policy designed to build capacity and increase output such as education and employment creation, and higher than average spending on negative areas such as criminal justice and unemployment benefit' (Taylor & Stanley 2005:xiii). The report ultimately resulted in a Human Rights and Equal Opportunity Commission case that was resolved in 2012 with significant additional funds provided for education in the community.

2.4 The wider Australian community

Accountability

In addition to Indigenous people, the wider Australian community has an ongoing interest in data about Indigenous people. In a democratic system, the population has an interest and obligation to ask whether a country's elected representatives are achieving what they have set out to achieve or claim to be achieving. To achieve this aim, data about Indigenous people is inevitably filtered by some of the groups discussed earlier. However, there is also a role for the media and statutory authorities to play in this accountability process.

An example of this accountability role is that provided by the Australian National Audit Office. During the 2011–12 and 2012–13 financial years, for example, the office used data on Indigenous people in such publications as: *Indigenous employment: the Australian Government's contribution to the Australian Employment Covenant; Capacity development for Indigenous service delivery; Implementation of the National Partnership Agreement on Remote Indigenous Housing in the Northern Territory; Indigenous Early Childhood Development; New directions: mothers and babies services; Indigenous secondary student accommodation initiatives; National Partnership Agreement on Remote Service Delivery; Indigenous employment in government service delivery; Indigenous protected area; and Australian Government coordination arrangements for Indigenous programs*. Many of these reports were critical of government practices.

Non-Indigenous people

The focus of Indigenous policy tends to be Indigenous people. Education disparities are to be reduced by encouraging their attendance, employment disparities by improving their skills, health disparities by reducing their alcohol and tobacco consumption, and so on. However, it needs to be recognised that the attitudes and behaviour of the non-Indigenous population are also important for improving Indigenous outcomes. Most of the teachers that Indigenous people interact with are likely to be non-Indigenous, as are most of the employers, service providers and health professionals they will encounter throughout their lives. Data for Indigenous Australians would therefore also consider the attitudes and behaviours of the non-Indigenous population, as well as how Indigenous outcomes compare with those of the non-Indigenous population.

Furthermore, there is an assumption that non-Indigenous ways of life are the norm to which Indigenous people should be aspiring to and moving towards. As argued in Walter & Andersen (2013:10), 'Not only are Indigenous people constituted as "the problem", non-Indigenous ways of life are left uncritiqued, despite the fact that in many cases current consumer lifestyles are environmentally unsustainable (to provide one example of many).'



Consider the case of discrimination in the labour market. There is certainly scope for discrimination of Indigenous people by Indigenous employers or co-workers (an example of which has been referred to as lateral violence). However, the attitudes of the non-Indigenous population towards the Indigenous population is likely to have a much bigger impact on whether the latter is able to gain and maintain employment and whether the remuneration received is appropriate to of the individual's skills and experience. Data from the Australian Reconciliation Barometer, conducted by Auspoll (2012) for Reconciliation Australia, is relevant here. Only a small minority of the wider Australian community reported favourable attitudes to Indigenous Australians in domains related to the labour market and education. That is, only 20% of the general community thought that Indigenous Australians were hard working, whereas 71% thought Australians in general were hard working. Furthermore, only 15% thought Indigenous Australians were disciplined compared to 41% for Australians in general.

Indigenous people are, of course, interested in the attitudes and behaviours of the non-Indigenous population. However, in terms of changing these attitudes and behaviours, it is the non-Indigenous population that will find this information most useful. Data like this are therefore important for understanding the experiences of Indigenous Australians in a range of domains.

3 The current statistical environment for data about Indigenous Australians

The previous section identified uses and users of data about and for Indigenous Australians. With a limited budget and, more importantly, a desire to limit the respondent burden on the Indigenous population, trade-offs will inevitably need to be made in terms of which uses take the highest priority. Furthermore, to ensure that as many legitimate needs as possible are met, the first priority needs to be making extensive use of the data that currently exist or that are to be collected anyway. It is only when gaps in the existing statistical environment are identified that new data collections should be considered.

There is a more extensive range of existing data sets and data sources about the Indigenous Australian population than there is for any other population subgroup of about the same size. The aim of this section is to provide an overview of the most important of these data sets, and to comment on their relative strengths and weaknesses.

This section discusses cross-sectional surveys (including the Census); key longitudinal data; qualitative data; administrative data; and finally, the promise and difficulties of community-controlled collections.

3.1 Cross-sectional surveys

A cross-sectional survey is one in which information is collected on individuals at one point in time only. Surveys can be repeated with individuals involved in more than one collection, but this information is not linked and individual outcomes are not tracked through time. There are 2 main types of cross-sectional surveys: (a) Censuses, where information is sought from all those in scope; and (b) sample surveys, where information is collected on a subset of those in scope. The results are then weighted to be representative of the population as a whole. In this section, I discuss 2 examples from each type.

Cross-sectional surveys can be used to identify the incidence or prevalence of a certain outcome and measure how such outcomes are changing through time if the data are collected consistently. They also allow researchers and policy makers to analyse the relationship between variables of interest.



Census of Population and Housing

The aim of the Census in Australia is to collect a range of demographic and socioeconomic information on the entire population of Australia (and the dwellings in which they live) at a particular point in time. The Census is far and away Australia's biggest data collection exercise, and unlike most other countries, it is carried out every 5 years. The most recent Census was carried out on 9 August 2011, and it collected data from everyone in Australia (excluding overseas visitors). As documented in the dictionary that accompanies each Census, for example, ABS (2011), information is collected about a range of topics, including demography, employment, education, need for assistance, language and unpaid work. Questions are about various geographic concepts, including a person's current place of usual residence, their usual residence 1 and 5 years ago, as well as their current place of work. This can be further combined with information on where a person was on the night of the Census to obtain a detailed geographic profile.

With some exceptions (as discussed below) the Census is a self-enumerated form with a single individual usually responding on behalf of all others in the household. On behalf of each person in the household, those filling out the form are asked 'Is the person of Aboriginal or Torres Strait Islander origin?' Options available are 'No', 'Yes, Aboriginal', and 'Yes, Torres Strait Islander'. A separate instruction states: 'For persons of both Aboriginal and Torres Strait Islander origin, mark both 'Yes' boxes'.

The 3 major advantages of the Census for analysis of Indigenous outcomes are the relatively consistent set of questions through time, number of Indigenous persons enumerated and the availability of a non-Indigenous comparison population. Specifically, there were 548,370 people who ticked 'Yes' to a least one of the Indigenous status options and 19,900,766 who ticked 'No'. Such large numbers allow for detailed analysis across a range of characteristics over very small geographic areas.

The main output from the Census is aggregate data or averages for different areas. These range in size from Statistical Area Level 1 (SA1), of which there were around 55,000 in 2011, to averages for states and territories or Australia as a whole. In addition to the aggregate data, the ABS also releases data at the individual level for a random 5% sample of individuals. This data set, known as the Census Sample File is available for analysis externally using the Remote Access Data Laboratory.

Despite these strengths, the Census does have some limitations. In addition to being cross-sectional (a point which will be returned to later), the fact that the Census is self-enumerated means that there is some uncertainty around the data provided: respondents usually are unable to immediately ask an interviewer for clarification on particular questions. The fact that it is self-enumerated also leads to a high degree of non-response. For example, there were almost twice as many records without a response to the Indigenous status question (1,058,583) as those who ticked 'Yes' to one of the Indigenous options.

One of the main uses of the Census, discussed in the previous section of this paper, is to accurately identify how many people there are (Indigenous and non-Indigenous) in Australia, different states and territories, and in particular localities. The self-enumerated nature of the Census makes this task somewhat difficult. Not only are there those that do not respond to the Indigenous status question (item nonresponse), but also a number of people who are missed from the Census entirely (unit nonresponse). The ABS attempts to mitigate this for the Indigenous population through the Indigenous Enumeration Strategy, which involves the use of interviewers in selected discrete Indigenous communities, as well as the use of Indigenous collectors and community engagement officers. Despite this, a large number of Indigenous Australians are missed from the Census with the ABS (2012) estimating that 17.2% of the Indigenous population who should have been counted in the 2011 Census was missed from the count. While this net undercount is adjusted for in the production of population estimates, those missed from the Census are likely to have different characteristics from those who are enumerated; hence socioeconomic and demographic analysis using the Census may be compromised.

A final limitation of the Census though is that it collects information only on a narrow range of data items. There is no information on a person's subjective views about their own life and, because it is targeted towards the total population, no information on Indigenous-specific notions of wellbeing. The only exception to this is



the important data item on language spoken at home, with 60,550 Indigenous Australians reporting that they spoke an Indigenous language at home in the most recent Census. Unfortunately, though, the Census has no information on the confidence the person has in that language.

Australian Early Development Index

A key policy question that has implications across a range of uses from Section 2 is whether Indigenous children start school with different strengths and weaknesses than their non-Indigenous counterparts. One set of data that has been constructed to consider this and similar questions is the Australian Early Development Index (AEDI). Collected for the first time in 2009 and repeated in 2012, the AEDI is based on a checklist completed by the teachers of children in their first year of full-time school. The checklist measures 5 key areas or domains of early childhood development: physical health and wellbeing, social competence, emotional maturity, language and cognitive skills (school-based), communication skills and general knowledge.

Like the Census, the AEDI is designed to be a population collection, with information sought about all children in their first year of full-time school. While coverage is not universal, information was collected for 261,203 children in 2009 and 289,973 children in 2012 (97.5% of the estimated national 5-year-old population in 2009 and 96.5% of children enrolled to start school in 2012). This high response rate can be attributed to the large amount of resources devoted to the AEDI and, in particular, the support given to the collection by teachers across Australia.

While it was not designed exclusively for Indigenous children, there was considerable effort devoted to ensure a large and representative Indigenous coverage rate. In total, there was information collected on 12,452 Indigenous children nationwide in 2009 and 15,490 in 2012.

The usefulness of the AEDI is limited by the fact that it is cross-sectional and selects one particular age cohort at one particular point in time. Furthermore, although considerable effort was devoted to ensuring the data collected were also useful in an Indigenous context, the measures of developmental vulnerability are not specific to the Indigenous population. Nevertheless, for the population in scope, it provides a very rich source of cross-sectional data.

National Aboriginal and Torres Strait Islander Social Survey and Australian Aboriginal and Torres Strait Islander Health Survey

In addition to the Census, the ABS runs a very extensive sample survey program in which information is collected on a subset of the population or a subset of particular populations. While these samples aren't always completely representative, considerable effort is made to make sure there is a sufficient sample from key population subgroups. Furthermore, weights are allocated to each person within the sample, which can then be used to make sure that estimates from the survey are applicable for the total in-scope population.

Many of the surveys that are carried out have a question that identifies whether the respondent is Aboriginal or Torres Strait Islander (or identifies as both). This includes the General Social Survey, the Labour Force Survey, the Survey of Education and Work, and the Survey of Disability, Ageing and Carers—all of which would potentially have information of relevance to governments, academics, Indigenous communities, other organisations and the wider Australian community. Because these surveys aren't targeted towards Indigenous Australians, they often do not have a sufficiently large sample to draw robust conclusions, and the Indigenous status of respondents is suppressed from the publicly available data.

Due to the limitations of the general survey program, the ABS also undertakes 2 very important Indigenous-specific surveys: NATSISS and AATSIHS. Recognising the limitations of the Census and other mainstream data collections, the NATSISS contains broad information across key areas of social concern for Indigenous Australians. Importantly, there were a number of wellbeing measures collected that were developed specifically for the Indigenous population, such as measures of cultural practices and the maintenance of Indigenous languages. The most recent NATSISS was carried out in 2008, with 2 previous surveys, one in 2002



and another in 1994 (when it was referred to as the National Aboriginal and Torres Strait Islander Survey or NATSIS). The next survey is due in 2014–15, and planning is well under way.

The Centre for Aboriginal Economic Policy Research has played a leading role in critiquing and analysing the 3 previous versions of the NATSISS/NATSIS (Altman & Taylor 1996; Hunter 2006 and Hunter & Biddle 2012). The strength of the NATSISS is that the information collected is of particular relevance to the Indigenous population. This is also a weakness in that it is not always possible to make comparisons with outcomes for the non-Indigenous population. For this, multiple data sources can be used. Furthermore, this is less of an issue when the General Social Survey coincides with the NATSISS (as occurred in 2002 and will again in 2014) but even then some comparisons are not possible due to differently worded questions.

The other strength of the NATSISS is that it collects information across a very wide range of topics. This allows for comparisons across a range of social, economic and health outcomes, with identified associations between variables often having high policy and research relevance. Once again, though, this could be seen as a weakness as there is limited information in each domain. Partly because of this, the ABS also surveys the health outcomes, behaviours and service usage of the Indigenous population—the AATSIHS. This survey (most recently carried out in 2012–13) is timed to more closely coincide with the health survey for the total Australian population—the Australian Health Survey (or in previous years the National Health Survey). Despite the AATSIHS and Australian Health Survey having considerable overlap in terms of timing and questionnaire content, the former still has Indigenous-specific questions.

Health is not the only domain for which a specially targeted Indigenous survey is needed. The Reconciliation Barometer mentioned earlier includes information from both an Indigenous sample and from the total Australian population on the relationship between the 2 populations and attitudes held by them. The National Indigenous Languages Survey contained a wealth of information on the use and strength of Indigenous languages and language speakers. Other gaps clearly remain though, and although respondent burden needs to be kept in mind, additional specific national collections could be considered in consultation with Indigenous people and organisations.

3.2 Longitudinal collections

Longitudinal data sets are sample surveys where individuals or households are surveyed more than once, with the information tracked through time. They are useful for analysing the factors associated with change through time for individuals and identifying possible causal relationships. This is a function that cannot be fulfilled by cross-sectional surveys.

There are 4 possible types of longitudinal collections, not all of which are available for the Indigenous population: cohort studies, panel studies, longitudinal administrative collections, and linked cross-sectional surveys.

Cohort studies follow a subset of the population who have some characteristic in common; more often than not, they share the same year or period of birth (also known as birth cohorts). There are a number of cohort studies of relevance to the Indigenous population (discussed below).

Panel studies select a representative sample of individuals or households and follow them through time. The most widely used panel study in Australia is the Household Income and Labour Dynamics in Australia (HILDA) survey. Unfortunately, the HILDA has only a small and unrepresentative Indigenous sample (it does not include remote Australia), and it can be used only for very broad Indigenous/non-Indigenous comparisons, with little capacity to look at variation within the Indigenous population.

While there are no other panel studies with a sufficiently large and representative Indigenous population (a point that will be picked up in Section 4 of this paper), there are some other longitudinal collections with important information about the Indigenous population. In particular, in late 2010 the Australian Government



commissioned the Melbourne Institute of Applied Economic and Social Research (at the University of Melbourne) to design and implement a new longitudinal survey, which was subsequently named *Journeys Home*. This survey tracked a national sample of individuals exposed to high levels of housing insecurity: it employed much more rigorous sampling methods than had ever previously been used. Importantly, the survey included a large sample of Indigenous Australians who were potentially exposed to homelessness.

There are also longitudinal administrative collections where clients of particular services or who receive particular payments are followed through time. These will be discussed in more detail in the next part of this section. The final type of longitudinal data is where 2 or more cross-sectional surveys are linked after processing has occurred using some form of identifying information.

Cohort studies—Footprints in Time and other collections

The most common type of longitudinal data set that is available to researchers and policy makers is cohort studies in which a sample of individuals that share a common set of characteristics (usually birth year or years) is followed through time. In Australia, the most widely used of these is the Longitudinal Study of Australian Children (LSAC), which is sometimes referred to as ‘Growing Up in Australia’. According to the data user guide for the survey, ‘LSAC aims to provide a database for a comprehensive understanding of children’s development in Australia’s current social, economic and cultural environment’ (AIFS 2011:8).

The LSAC was constructed around 2 cohorts—the B cohort (born March 2003–February 2004) and the K cohort (born March 1999–February 2000). In Wave 1 of the LSAC, there were 4,983 children aged 4–5 years, of which 187 were identified as being Indigenous (3.8% of the sample). By Wave 4, when the children were aged 10–11 years, there were 3,940 children left in the LSAC, of which 105 (2.7%) were identified as being Indigenous. Clearly this is a very small sample of Indigenous children and does not allow for detailed analysis within the population. Furthermore, sample attrition (or dropping out of the survey between waves) was much higher for the Indigenous population than for the non-Indigenous population.

It is not yet possible to use the LSAC to look at school outcomes and transition into the labour market. An alternative set of data, the Longitudinal Surveys of Australian Youth (LSAY), focuses on youth outcomes and transitions by following successive cohorts of 15-year olds. Furthermore, unlike the LSAC, the 2006 and 2009 cohorts of the LSAY have a large and much more representative Indigenous sample. While the data for 2012 are not yet available, Wave 1 of the 2006 cohort includes information on 14,170 respondents who were aged about 15 years at the time of the survey. Of these, 1,080 were Indigenous, with 42.3% of the Indigenous sample attending a school in a major city, 46.7% attending a school in provincial Australia, and the remaining 11.0% attending a school in remote Australia. The 2009 LSAY had information on 14,251 children, who were also aged 15 years at the time of the survey. The Indigenous sample in 2009 was even larger than in 2006, with 1,143 respondents: 48.9% of the Indigenous sample was attending a school in a major city, 42.3% was attending a school in provincial Australia, and the remaining 8.8% was attending a school in remote Australia.

The limitation of the LSAY is that sample attrition is even higher than in the LSAC. Only 518 or 48.0% of Indigenous students who were in Wave 1 of the 2006 survey remained in the survey for Wave 2, whereas 67.5% of non-Indigenous respondents remained. Attrition rates decline in subsequent waves, with 84.4% of the total population surveyed in Wave 2 being available in Wave 3, and 79.9% of these respondents being available in Wave 4. These high rates of attrition are particularly problematic if those who remain in the sample have different characteristics to those who drop out.

Partly in response to the low sample of Indigenous children in the LSAC and the reasonably high rate of sample attrition, considerable resources have been devoted to the development of an Indigenous-specific cohort study—the Longitudinal Study of Indigenous Children (LSIC). Also known as *Footprints in Time*, the LSIC is the first large-scale longitudinal survey in Australia to focus on the development of Indigenous children. The first wave



of the survey was carried out between April 2008 and February 2009, and it collected information about 1,687 study children and their families.

Like the LSAC, the sample for the LSIC was designed around 2 cohorts—babies (born between December 2006 and November 2007) and children (born between December 2003 and November 2004). The eventual sample comprised of 960 children in the baby cohort and 727 in the child cohort. While the survey administrators aimed to keep the sample within these birth date ranges, in practice there was a minority of children in the sample who fell outside them.

According to the Commonwealth department that administered the LSIC at the time, the main objective of the study is to provide high quality quantitative and qualitative data that can be used to provide a better insight into how a child's early years affect their development. Specifically, *Footprints in Time* has 4 key research questions:

- What do Aboriginal and Torres Strait Islander children need to have the best start in life to grow up strong?
- What helps Aboriginal and Torres Strait Islander children to stay on track or get them to become healthier, more positive and strong?
- How are Aboriginal and Torres Strait Islander children raised?
- What is the importance of family, extended family and community in the early years of life and when growing up? (FaHCSIA 2009).

By focusing on the specific needs and circumstances of the Indigenous population, there has been considerable support within the community for the aims and goals of the LSIC. This is seen in the relatively high sample retention achieved across the first 4 waves of the survey. Ignoring those who were added to the sample subsequent to Wave 1, 1,031 out of 1,671 children (61.7% of the original sample) have been counted in all 4 waves (FaHCSIA 2013:65). While enough children have dropped out of the sample to worry about selective attrition, this is a reasonably high retention rate given the mobility of the population in question. However, it has come at a cost. There is a trade-off made between community support and representativeness on the one hand and the need to control for geographic clustering of individuals and the associated effect on standard errors and inference on the other hand (Hewitt 2012).

Linked cross-sectional data sets

An alternative to following individuals through time is to link 2 cross-sectional data sets that, in theory at least, contain significant overlap in the population. This approach is cheaper than longitudinal data and it avoids the problem of attrition that longitudinal data suffer from.

This linking is ideally done by using some form of time invariant information that is unique to each individual. With administrative data, this might be, for example, a person's customer reference number. When that is not available, other information that might be used includes a person's name, their date of birth, or where they lived at a particular point in time.

A new set of data that takes the latter approach is the Australian Census Longitudinal Dataset (ACL), which was released by the ABS in late 2013. According to the ABS (2013c), 'a sample of almost one million records from the 2006 Census (Wave 1) was brought together with corresponding records from the 2011 Census (Wave 2) to form the largest longitudinal data set in Australia.' In essence, 5% of records from the 2006 Census are linked probabilistically based on the most likely match given observed characteristics with available data from the 2011 Census. As this linking was done without the aid of the individual's exact name and address, there will be a minority of linked pairs that are not in reality the same individual. This needs to be kept in mind when making conclusions based on the data.

There are limitations of the ACL. There is an under-representation of the Indigenous population, which was foreseeable and could potentially have been negated by an oversampling within the original 5% sample.



Furthermore, the use of Indigenous status as one of the variables in the linking process limits (though not negates) what we can say about those whose identification changes through time. Nonetheless, the ACLD will allow for the analysis of change through time in standard Census measures of demographic, socioeconomic and geographic characteristics. Perhaps more importantly though, for the first time in Australia, we have a large data set with information on a person's Indigenous status in both 2006 and 2011. This will potentially allow for analysis of the factors associated with identification of Indigenous status, a topic of relevance for government, academics and community organisations.

3.3 Qualitative data

The focus of this paper up until now has been on quantitative data sets. That is, those that attempt to collect information on the same topic in a systematic way across samples that are large enough to be representative of the population of interest. It should be noted though there are a number of policy-relevant uses of qualitative data about Indigenous Australians that should be touched on briefly. Rather than numerical constructs, qualitative data involve information such as words, pictures or objects that aim to gather an in-depth understanding of human behaviour and the drivers behind such behaviour. Furthermore, qualitative data are more explicitly situated in a particular context, time and place with less of a claim to universality or high level abstraction.

There is a range of techniques for the collection and analysis of qualitative data. Three stand out as being of particular relevance for many of the uses of data discussed in Section 2 of this paper. The most in-depth is participant observation, where researchers typically interact very closely with members of a culture, group, or setting, potentially adopting roles to conform to that setting. The aim of this type of data collection is to gain a closer insight into the culture's practices, motivations and emotions. A second approach, which has a slightly greater distance between the researcher and subject, is the interview approach. This could be through a structured interview, in which each interviewee is presented with exactly the same questions in the same order, or a semi-structured or unstructured interview, in which new ideas can be brought up during the interview as a result of what the interviewee says. The third main approach is focus groups, in which questions are asked in an interactive group setting with participants free to talk with other group members.

In parallel to the critique of quantitative research methods by Walter & Andersen (2013) and Yu (2011), there is an extensive body of research on Indigenous research methodology from either a qualitative perspective or that touch on qualitative data. Highly cited and influential are *Decolonizing Methodologies: Research and Indigenous Peoples* (Smith 1999), as well as Karen Martin (Booran Mirraboopa's) (2003) article in the *Journal of Australian Studies* on 'Ways of knowing, being and doing'. While there is no definitive set of Indigenous research methods, most involve Indigenous people throughout the research process, with agreement on questions and methodological approaches and shared ownership of data and knowledge. As argued in Walter & Andersen (2013) though, these principles of shared ownership are as relevant for quantitative as qualitative data.

In many ways, qualitative data can be seen as a complement rather than a substitute for quantitative data. Indeed, a whole issues paper could be devoted to the topic because qualitative data provide a rich description of a phenomenon, including why and how people behave and make decisions. Furthermore, qualitative data can be useful for generating hypotheses and building theories, which can then be tested using alternative data sets. Qualitative data can give insight into how a policy or program is actually implemented on the ground, how participants understand a policy and program, and how this might be different for individuals in specific cultural contexts. In terms of policy though, qualitative data are less useful for analysing the geographic or socioeconomic distribution of need or testing the effectiveness of a particular program because the conclusions are based on the specific context and are not always generalisable.



3.4 Administrative data

A lot of data are generated from the normal day-to-day running of government programs. Clients are kept track of, their outcomes measured, and occasionally, demographic and socioeconomic data are collected as well. Most programs also collect information on whether their clients identify as being Aboriginal or Torres Strait Islander, with that information also retained. Increasingly, such service usage data is being used and analysed for a much greater range of purposes than they were originally collected for. This creates privacy concerns with the need for safeguards to ensure service users aren't identified as part of this research. In addition, when programs or policies are trialed and evaluated, a large amount of data is collected and analysed. This part of the paper looks at the types and availability of service usage and trial data.

Service usage data

As computing power and the ability to store data continues to increase, governments and other service providers are willing and able to retain a much greater range of data generated by the day-to-day running of government programs. Data has always been collected on whether or not an enrolled child is attending school on a particular day, when someone sees their local doctor, when they access welfare payments, and so on. What has changed, however, is the ability to store and analyse that data, as well as the consistency with which information on whether the user of a service is Aboriginal or Torres Strait Islander is collected and recorded.

In 2010, the AIHW published a paper on *National best practice guidelines for collecting Indigenous status in health data sets*. In it, the AIHW documented the standard Indigenous status question (in essence the one included in the most recent Censuses) as well as standard practice for how to ask the question and how to record the response. While the focus of AIHW (2010) is on health data sets, the principles hold in many other contexts. The *Report on government services 2013: Indigenous compendium* (SCRGSP 2013) provides a wealth of information from such data sets, albeit at the aggregate rather than individual level.

Service usage data generated from people's response to this question can be used for a number of purposes. Most obviously, it can be used to identify how many (Indigenous and non-Indigenous) people are using a particular service at a particular point in time and how this might be changing through time. If a suitable denominator is available, then these numbers can be turned into rates, which would allow comparisons to be made within the Indigenous population and between Indigenous and non-Indigenous Australians. Service usage data, when linked with other information, can also be used to identify the outcomes of those who use a particular service. In some cases, the outcome is reasonably straightforward to measure. In other cases, the outcome of the service needs to be linked with other data held by that agency or another agency. For example, linking school attendance data to school completion data and program data can tell us a lot about the effectiveness of particular school programs. It should be noted, though, that the outcomes available on administrative data do not always capture the broader notions of wellbeing that policy is ultimately aiming to improve.

While useful, there are other limitations of service usage data for understanding the outcomes of the Indigenous population and the effectiveness of government programs in improving Indigenous wellbeing. From a quality point of view, there is still variation in the extent to which the Indigenous status question is applied consistently either at the program level or by individual staff of service providers. According to AIHW (2012:vii), 'Although a number of agencies have adopted the ABS standard ... there are still inconsistencies in the collection and recording practices used for a number of data collections.' This can lead to either a high rate of non-responses to the Indigenous status question, or uncertainty around whether those identified as Indigenous or non-Indigenous would in fact normally identify as such.

A related issue to Indigenous identification is that it is sometimes difficult to identify a suitable denominator for administrative data from which rates can be calculated. This is not always the case. For example, when calculating attendance rates, the denominator is all those enrolled in a particular school or jurisdiction. In other cases though, the denominator is the entire population. For example, when calculating the percentage of the



Indigenous population enrolled in preschool, the denominator is the total Indigenous population aged 4. While an estimate for this is available, it might not be consistent between Census years. Furthermore, patterns of identification in the Census might be different to patterns of identification in the relevant administrative data sets. This can lead to misleading conclusions regarding relative rates or change through time.

In terms of analysing the effectiveness of particular programs using service usage data, quite often information on the comparison group is not available, as by definition they are not using the service. This problem can be mitigated if there are other sources of information, for example survey data or data from other services. However, this is not always available and when it is there is no guarantee that the data and outcomes are comparable. Perhaps more importantly, without being able to control for non-random selection into the program, it is usually not possible to tell whether those using other services or no services at all are otherwise comparable. The reasons for selection into the program are often predictors of the outcomes of that program.

A final limitation of service usage data is that they are not always, and some might argue rarely, available for access by external users (at least at the individual level). This can mean that it is difficult to link data sets across important domains. More importantly, those who are analysing the data are often in the same organisation that is providing the service: this is a potential conflict of interest when it is the quality and effectiveness of that service which needs to be evaluated.

There are health-related data collections that involve the recording of data at the Indigenous-specific primary health care level, where individual organisations (community-controlled and non-community-controlled) report against particular indicators and aggregate data are publicly released. As well as this, unpublished individual reports are provided to each service to contribute to their quality improvement processes and monitor their progress overall. This includes the Aboriginal and Torres Strait Islander health services Online Services Report (AIHW 2013a) and the National Key Performance Indicators for Aboriginal and Torres Strait Islander primary health care (see Box 2).

Box 2: National key performance indicators for Aboriginal and Torres Strait Islander primary health care

A set of national key performance indicators for Indigenous-specific primary health care services were developed by the former Department of Health and Ageing (now the Department of Health), state and territory health departments, and the AIHW. The indicators are designed to monitor government-funded primary health care services provided to Aboriginal and Torres Strait Islander Australians, and to monitor progress towards the COAG Closing the Gap targets—in particular, the targets for life expectancy and child mortality.

In 2012–13, the AIHW began to collect data for the first 11 of the 24 indicators, and an individual report containing data analyses and interpretations was provided to each service to help with their continuous quality improvement activities, staff training, and presentations of aggregate information about their clients. It is anticipated that the remaining indicators will be introduced progressively, as they are approved and incorporated into the data collection process (AIHW 2014).

Trial data

The service usage data discussed above often has very useful information about those who are using a particular service and what some of their outcomes are. This information allows for the needs of current clients to be better met and the costs and implementation issues identified.

Crucially though, much of the data on programs for Indigenous Australians do not tell us much about those who are not using the service or what the outcomes of those who did use the service would be if they chose not to do so. We are often missing the counterfactual. To draw inferences on specific policies, therefore, information on relevant outcomes would need to be integrated with the types of randomised policy trials advocated by Leigh (2010). To consider how such trials would work, it is useful to step through a stylised example of a policy evaluation. The first step would be to identify the target population for a particular policy. As an example, this



might be school-aged students at risk of dropping out. The next step would be to randomly assign a subset of individuals from that target population into a treatment and a control group. This assignment could be by school, by class within school, or by geographic area. The important point is that assignment is done randomly so as to minimise the chance that the treatment and control group have different, unobservable characteristics.

Having identified the treatment and control group, the next step would be to apply the policy intervention to the treatment group. The intervention could be, for example, a conditional cash transfer (that is, additional payment to the family conditional on attendance), an extra Indigenous teacher in the school, modification to the curriculum, or a voucher to attend a boarding school. For the purposes of this paper, the particular policy doesn't matter. What matters is that the control group doesn't receive the same intervention as the treatment group, or receives it at a later stage. The latter approach is known as a pipeline comparison (Ravallion 2008) and is useful if those designing the evaluation would like to avoid those in the control group missing out on the intervention entirely.

The final stage in the policy evaluation is to test for differences in outcomes between the treatment and the control group. If the treatment and control group are defined well (that is, randomly without any bias), it is not necessary to take a baseline measurement as one can assume that they do not differ substantially. However, baseline comparisons can be useful to test for observable differences between the 2 groups. The particular outcome measure analysed will be influenced by the policy being evaluated. In the example outlined above, one of the outcomes measured would be school completion and drop-out rates.

There are arguments against such randomised controlled trials. The first of these is that it is unethical, in that governments are potentially withholding beneficial programs from individuals based on, in effect, the toss of a coin. However, the point is that we don't actually know that it is a beneficial program. That is why we are evaluating it. A second argument is that such evaluations are costly, as providers of programs have to spend money on follow-up surveys. This is certainly true from the point of view of individual program deliverers or service providers, and the costs should be supported in funding allocations. However, from the perspective of society as a whole, spending money on evaluating programs properly is far more cost-effective than spending money on ineffective policies.

A final issue is political. Backers of policies don't necessarily want it to be known that their policies were ineffective. This point was made well in a quote from Campbell (1969), which was quoted in James (2013:108). Specifically, 'If the political and administrative system has committed itself in advance to the correctness and efficiency of its reforms, it cannot tolerate learning of failure.' This is by no means a blanket statement, and there are of course many instances of governments changing their policy focus as new evidence or evaluations come to hand. A more important point though is that independently evaluated programs found to be successful are a lot more difficult to remove when government changes.

There are, of course, a number of quite valid limitations to randomised controlled trials. First, withdrawing treatment for the treatment or control group could have large negative effects. For this reason, one would be reluctant to test the effect of removing income support or other essential social services. A second limitation of randomised controlled trials is that they are often possible for only small programs that are not always scalable. Related to this, it is not always possible to measure spill-over effects or the impact of the intervention on those not in the treatment group. Randomised controlled trials are of less use for setting broad policy direction, and this is where evidence from surveys like the ones discussed earlier will always be useful. An additional limitation is that quantitative evaluations like those outlined above do not measure all of the potential costs and benefits of a program, and in particular, they are not very good at testing for unforeseen outcomes. In that respect, they should ideally be integrated with qualitative evaluations (on a sample of the treatment and control groups if needs be) of the type outlined in Rao and Woolcock (2003).

Cobb-Clark (2013) identified a number of other relevant Indigenous-specific limitations of randomised controlled trials. Indigenous people make up a small percentage of the population, so sample sizes in standard surveys are small. It is also often difficult to define a relevant control group because of the potential heterogeneity across communities. Furthermore, given the range of interventions being undertaken in many communities, it is often



difficult to separately identify the effect of one intervention or program over the other. Most importantly though, and this is a tension with other desirable aspects of Indigenous data, interventions that need the approval and support of community members compromise the random element of the design. As Cobb-Clark (2013:86) argues, 'The effects of the selection process itself—normally long, drawn-out negotiations between government and Indigenous Elders—will be a component of what is measured in the estimated treatment impact.' There are good reasons for this negotiation, some of which are covered in this paper. However, it needs to be kept in mind that it comes with a trade-off.

One other point made by Cobb-Clark that is crucially relevant for this paper is the need for trial data to be made available to external researchers and that evaluation be done independently from government. The author argues on page 90 that 'The lack of a willingness to commit to eventual publication of results has meant that Australian academics are increasingly disengaged from evaluations of major economic and social initiatives' and on page 91 that 'all components of any program evaluation, including the unit-record data on which it rests, must be widely and publically available, so that results can be replicated and confirmed'.

None of this discussion should be taken to mean that there are no useful evaluations or use of trial data in Indigenous policy (CtGC 2013). However, it would be fair to say that the majority of the evaluations that do occur focus on implementation rather than outcomes. For example, consider the Sentinel Sites Evaluation of the ICDP. This is an innovative, formative evaluation of 24 sites across Australia by the Menzies School of Health Research with funding from the former Commonwealth Department of Health and Ageing. However, the main achievements from the implementation of the ICDP are listed as follows:

- improved access to primary health care services and to affordable medication for many Aboriginal and Torres Strait Islander people
- improved orientation of the general practice sector to the needs of Aboriginal and Torres Strait Islander people
- significant steps towards the early establishment of a new workforce that is focused on health promotion and in development of local health promotion initiatives
- increased attention to enhancing access to specialist, allied health and team-based approaches to chronic illness care (Bailie et al. 2013:10).

These are, no doubt, important findings for the implementation of a particular program. However, they tell us very little about the impact of the ICDP on rates of chronic disease due to the time lag between when a program is implemented and when changes in outcomes could be observed. This is in part due to the long time lags involved but also because of the difficulty in attributing effects to a particular intervention in a context of many competing policies and programs.

Outside of the health domain, there are even fewer evaluations. Ultimately, one of the main limitations of data for and about Indigenous Australians is the lack of trial and program data that will enable the evaluation of the effectiveness of programs and policies.

3.5 Community-controlled collections

Collections that are designed, funded and administered by communities themselves can provide an alternative source of data with a much tighter focus on the needs of the Indigenous community. These community-controlled collections have much greater scope than other approaches to focus on the type of information that is of relevance to those who are providing the data. This increases the chances of a more holistic understanding of Indigenous development and a chance to counterbalance the deficit- and gaps-approach taken to much of the data collection and analysis. While relatively new, community-controlled collections are providing information for policy and community discussions at the local level.



This approach contrasts starkly with that discussed earlier in this section. For example, because the Census is a national-level data set with a particular aim and purpose, it has very little information that is specific to the needs of the Indigenous population, apart from the question on Indigenous status. So, while it is possible to analyse the results at the small area level using the Census, this can only be done using variables that do not necessarily reflect the needs and aspirations of the Indigenous population. Furthermore, most administrative data sets are collected to meet the needs of service providers, rather than the Indigenous population itself.

There is much greater scope for input into the design and collection of Indigenous-specific data sets like the NATSISS and the LSIC, with the latter, in particular, having a very extensive consultation process, support from the communities involved, and the capacity to incorporate the skills and knowledge of Indigenous interviewers (FaHCSIA 2009). However, these sample surveys do not have sufficient sample size to provide information for individual locations or individual communities.

The absence of Indigenous specific information in a number of those collections is both a cause and a consequence of the fact that there are very few Indigenous people taking a leading role in the collection and analysis of large-scale survey data in Australia. There are of course exceptions, including the National Advisory Group on Aboriginal and Torres Strait Islander Health Information and Data, which acts as a means of bringing together Indigenous people and organisations, academic researchers and government agencies involved in the collection and analysis of data at a national level. However, Walter & Andersen (2013), both Indigenous researchers in academia, argue that:

Although Indigenous people are the subject of Census and administrative collections around the health, welfare and justice systems, it is largely non-Indigenous researchers undertaking research using this data.

and furthermore that:

This lack of an established Indigenous presence, combined with the specific and technical language used and the statistical basis of quantitative analysis create an atmosphere around the practice that is alien to many Indigenous researchers. Walter & Andersen (2013:133).

There are limitations to community-controlled data sets. By definition they are specific to the community in which they take place and are therefore not useful for comparative purposes. It should be noted though that this is not what they are designed for. More importantly, though, they are expensive and require considerable financial and in-kind support from the community. The Yawuru 'Knowing Our Community' (YKC) survey (Box 3) was successful because of the Native Title settlement that preceded it and, it could be argued, because of partnerships with external organisations that had significant statistical expertise and experience. This might not be necessary for future surveys in Broome due to the experience gained from the YKC survey. However, other communities that would like to have similar information about themselves could lack the required resources and expertise. It is for this reason that Taylor (2013:127) has argued that 'what they [groups such as the Yawuru] seek from government is ... support for capacity building in their compilation and use of customised data as a means of promoting their full and effective participation in local governance and development planning.'

Box 3: The Yawuru 'Knowing Our Community' survey

An example of a community-controlled collection is the Yawuru 'Knowing Our Community' (YKC) survey that took place in 2011 in Broome in the northwest of Western Australia.

One of the aims of the survey was to collect an accurate population baseline against which data from the relevant Census could be benchmarked. However, it had 2 additional aims in that it attempted to obtain an estimate of the Yawuru (as opposed to Indigenous) population of Broome, as well as to gain information of particular relevance to the community.

According to Taylor et al. (2012:8) it was 'the first survey to be knowingly comprehensive in coverage and to be developed, managed, conducted and controlled entirely by local Indigenous organisations and local Indigenous residents for the primary purpose of informing their own local planning needs'.



As mentioned earlier, data are also collected by community-controlled health services as part of reporting processes to government. These data can be used within the health services for their own quality improvement needs, and are also publicly released at an aggregate level for use by a wider audience. As well, in 2013, the AIHW and the National Aboriginal Community Controlled Health Organisation publicly released a report card on the Healthy for Life health outcome data from a subset of the Indigenous community-controlled health services involved in the Healthy for Life program (AIHW 2013b).

4 Gaps between data needs and data availability

The previous 2 sections detailed a range of uses of data about and for Indigenous Australians as well as sources of such data. While there are limitations in the latter, it should be kept in mind that the body of empirical data in Australia is second-to-none internationally when it comes to a country's Indigenous population. While New Zealand arguably has a greater range of Māori-controlled collections and evaluations, Canada and the United States, which have Indigenous populations that make up a similar share of the population to Australia, do not have the same range of Indigenous-specific, nationally representative surveys. Despite this, there are still gaps between data needs and data availability. Furthermore, as will be discussed in this section, it is quite clear that there is no single data set that meets all the needs of government, academics and researchers, Indigenous communities, other organisations, and the wider Australian community.

The following table summarises the characteristics and potential uses of 13 of the most commonly used data sets or types of data sets about or for Indigenous Australians. Those data sets that have a particular characteristic or can be used for a particular use are marked with an 'x'. When it is a little more ambiguous, an 'o' is used. Blank cells are for when the data set definitely does not have that characteristic or cannot be used for that purpose.

Table 1: Sources of data about Indigenous Australians and their uses

Characteristics/uses	Cross-sectional data				Longitudinal data				Qualitative	Admin data		Community controlled
	Census	AEDI	NATSISS	AATSIHS	LSAC	LSAY	LSIC	HILDA		ACLD	Service data	
Characteristics												
Large Indigenous sample	x	x	x	x	o	o	x	x	x	x	o	x
Nationally representative sample	o	x	x	x		o	o	o				
Non-Indigenous comparison	x	x		o	x	x	x	x	x	x	o	
Information across life course	o		x	x			o	o				o
Indigenous-specific measures			x	o			x		x		o	x
Uses												
Government/policy makers												
Allocation	x									o		
Need	x	x	o	o		o	o	o		x		
Monitoring	o	o	x	x						x		x
Evaluation					o	o	o	o	o	o	o	
Planning	x							x				x
Academics/other researchers												
Understanding	o	x	x	x	o	o	x	o	x		x	x
Explaining	o	o	x	x	o	o	x	o			x	
Indigenous community												
Reflection			o	o			x		x			x
Advocacy	x	x	o	o		o	o			o		x
Wider Australian community												
Accountability	x	x	x	x				x		o	x	o
Non-Indigenous people							o		x		o	

It is clear from Table 1 that there is no data set that has all the desirable characteristics in the first part of the table or can be used for all the purposes discussed in Section 2 of this paper. Perhaps more so than in many other policy domains, a range of data is required to ensure policy settings will meet the needs and aspirations of Indigenous Australians.



Focusing on the characteristics of the data sets to start with, it is clear that there are a number of data sets that have large Indigenous samples. Two of the longitudinal data sets (the LSAY and the LSAC) are marked with an 'o' due to the high rate of sample attrition. That is, they start off high, but decline through time. In terms of being nationally representative, the AEDI and 2 ABS surveys best meet this aim. The Census (and ACLD) are marked with an 'o' due to the high rates of non-response discussed earlier. While the large Indigenous samples are positive in terms of the ability to analyse and understand the Indigenous population, it does highlight the very large respondent burden placed on the Indigenous population. Any additional data collections would need to recognise and minimise this.

One limitation of the data sets listed in the first part of Table 1 is that there is no longitudinal data set that has a large Indigenous sample and information across the life course. The ACLD does to a certain extent, but there is very little information in the Census on child health and development outcomes. Furthermore, while potentially promising, the ACLD has not been evaluated for the robustness of the linking with regards to the Indigenous population.

Looking at the second part of Table 1, the ABS survey program (which includes the Census, NATSISS, AATSIHS and ACLD) is clearly being used to support the needs of government and policy makers. For allocation, need and monitoring, service data is also an important source of information. With so many collections being used, however, this raises the issue of the importance of ensuring definitions and scope line up. There is also the issue of the timing of the surveys. For allocation and planning in particular, we are reliant on the Census and the ACLD. However, these are 5-yearly collections with a significant temporal lag between the collection and output of data. This would be less of an issue with a stable population (and is much more frequent than in many other jurisdictions), but with a rapidly changing population, this can be problematic.

The biggest gap for government and policy makers, however, is in terms of evaluation. Some of the longitudinal data can, at times, be useful for evaluation, and administrative data can also potentially be useful. These data sets are useful, though, only if the relevant policies and programs are designed in such a way that they can be evaluated using the data. In this case, it is not always the data that are limited, but the interaction between the data and the programs.

It would appear from the table that the needs of academics and researchers are well met by the current statistical environment. There are a number of data sets that can be used, many of which have a large Indigenous sample, a non-Indigenous comparison, information across the life course and Indigenous-specific measures (though not all 4 simultaneously, it should be noted). What is not mentioned in the table though, is some of the difficulties in accessing that data. Unpublished administrative data and in particular unit record data is often very hard to obtain and none of the ABS collections are available for download or order. Analysis of the surveys is either done through request to the ABS for specific cross-tabulations or through submitting code to the Remote Access Data Laboratory. These restrictions are present to ensure that the confidentiality of the data is maintained.

While confidentiality is clearly an important consideration for the ABS, one should also recognise that it does place restrictions on the types of analysis that can be carried out and the level of geography that can be used. Despite there being 3 different collections of the 2008 NATSISS available, none of them allows for analysis of variation within remote Australia (that is, differences between remote and very remote areas). Nor do they allow for analysis of variation within non-remote Australia (that is, comparisons between city and regional areas). In addition, the geography for the Census Sample File is constructed to suit analysis for the total population with no ability to compare outcomes within the Northern Territory (for example). ABS could consider facilitating researcher access to a much greater range of Census and survey data without risk to confidentiality for individuals.

Indigenous community and other organisations are less well served than governments or academics by the extant Indigenous data. As argued in Yu (2011), Malezer (2013) and Taylor (2013), the existing data are not available at the level of geography that is needed by Indigenous organisations, nor is the type of data that are available always suitable for reflecting on progress in the areas that are meaningful for the Indigenous population. This puts a heavy burden on community-controlled collections that, while an important future source of data in Australia, are not always easy for all communities or people to collect.



5 Concluding comments and how Indigenous data could be improved

The aim of this paper was to provide a summary of the current strengths and weaknesses of data about and for the Aboriginal and Torres Strait Islander population of Australia. A range of uses for such data was identified, with different users having a particular focus. These users and uses are summarised below.

- Government and other policy makers
 - Allocation: how should we distribute funds to jurisdictions and other service providers?
 - Need: where are the areas of absolute and relative need?
 - Monitoring: is progress being made in meeting targets at the national and sub-national level?
 - Evaluation: how effective are particular programs and interventions in achieving their aims?
 - Planning: what is the likely structure of the future Indigenous population?
- Academics and researchers
 - Understanding: what are some of the patterns and diversity within the Indigenous population?
 - Explaining: what are the determinants and correlates of relevant outcome variables?
- Indigenous community and other organisations
 - Reflection: how is a particular community of Indigenous people progressing?
 - Advocacy: are particular communities or people receiving the level of government support that they are entitled to?
- The wider Australian community
 - Accountability: are our elected representatives achieving what they have set out to achieve or claim to be achieving?
 - Non-Indigenous people: what are the general attitudes and behaviours of the non-Indigenous population, and how do outcomes compare with the non-Indigenous population?

It is clear from the discussion in this paper that governments and academics and researchers are generally well covered by the available data, but there are still a few key gaps. Moreover, the needs of Indigenous community and other organisations are not as well met, with much of the burden of providing data for these users placed on community-controlled collections. Other gaps and issues include:

- complexities around identification, where the same individual could be recorded as having a different Indigenous status across collections or through time
- no data set with longitudinal information on Indigenous Australians across the life course (that is, from childhood through to adulthood)
- the use of a range of data sets for a single purpose, with the result that the scope and definitions might not be consistent
- the 5-year gap between each Australian Census, a data set that is the only source of data for some of the uses discussed above (in particular, allocation and planning)
- difficulties in accessing administrative and other data from outside government
- lack of data at the right level of geography, for the right population and for the right concepts to be useful for Indigenous community and other organisations
- the lack of trial and program data that will enable the evaluation of the effectiveness of programs and policies.



With these limitations in mind, there are 4 suggestions for improving data about and for Indigenous Australians.

It is worth emphasising first, though, that like with most other policy domains, there are inherent trade-offs with Indigenous data. Some have already been discussed. For example, community input into the design and delivery of government programs is fundamental to meeting the needs of Indigenous people and a key aspect of the United Nations Declaration on the Rights of Indigenous Peoples (which Australia is a signatory to). However, such community input makes it very difficult to do the type of evaluations that governments need and communities demand in order to know whether a particular policy is working or not. These trade-offs can't always be mitigated, but neither can they be ignored.

Indigenous specific information in the Census

The main aim of the Census is to count the number of people in Australia to allocate seats in Parliament and funds across jurisdictions. It is also used to understand how the broader Australian society is changing through time. For this reason, it is not possible to structure it in a way that focuses exclusively on the needs of the Indigenous population. However, there is some scope to change the way in which data are outputted in terms of geography and also to make sure categories and the cut-offs used (for example with regards to income, number of bedrooms and number of children ever born) reflect Indigenous aspirations and circumstances.

As the Census moves to being filled out either online or, in the case of discrete Indigenous communities, by interviewers, there is scope to include questions in the Census that are specific to a particular group of the population and that are skipped by the rest of the population. The potential for such Indigenous-specific information was discussed at a workshop in November 2013 at the Sydney Centre for Aboriginal and Torres Strait Islander Statistics with a number of options discussed (for more information see:

<http://sydney.edu.au/health-sciences/scatsis/events.shtml>). For example, if a person answers that they are either Aboriginal or Torres Strait Islander, then this might prompt them to be asked a few subsequent questions. This might relate to identification within a particular Indigenous group, or it might relate to attachment to a particular geographic area (which might be different from where they currently live).

Another additional set of questions that might be considered for the Census is on language usage. It was noted before that the only Indigenous-specific measure on the Census apart from the identification question is whether or not a person speaks an Indigenous language at home. However, there is no information on the extent of the person's knowledge, whether they speak the language outside the home, or whether they identify with a particular language group. Following the recommendation from AIATSIS & FATSIL (2005:119), the ABS should consider including such information for those who respond yes to the Indigenous status question.

Longitudinal data

One of the limitations mentioned above is the lack of longitudinal data with Indigenous-specific information across the life course. This is arguably one of the biggest gaps in terms of the type of information that would allow for a rigorous understanding of the determinants of Indigenous socioeconomic change and improvements in Indigenous and non-Indigenous notions of wellbeing; but longitudinal databases are expensive. They also place a considerable burden on respondents since Indigenous Australians already provide a range of information to government statistical agencies.

Biddle and Yap (2010a) outlined an alternative proposal that would integrate existing ABS surveys in a way that would provide such information but that would not either be too expensive or significantly increase respondent burden. In what we labelled the National Closing the Gap Survey (NCGS) a rolling-panel approach to the collection of national statistical data sets on Indigenous Australians could be implemented. Under this approach, households or individuals are retained in the sample for a fixed number of surveys but are eventually dropped out and replaced to keep the data representative of the nation as a whole. A hypothetical structure of a 6-year collection cycle might begin with a NATSISS in 2014, another in 2020, as well as an AATSIHS in 2017 (6 years after the next survey) and 2023. In the intervening years, Biddle and Yap (2010a) proposed that a reduced module of



questions be asked that allow key measures of wellbeing and the government’s Closing the Gap targets to be analysed. Depending on costs, the NCGS could be carried out on a subset of the original cohort only. A proposed structure is outlined in Table 2.

Table 2: Proposed structure of the National Closing the Gap Survey (NCGS)

Year	Cohort 1	Cohort 2	Cohort 3	Cohort 4
2014	NATSISS/General Social Survey			
2015	NCGS			
2016	NCGS			
2017	AATSIHS	AATSIHS		
2018		NCGS		
2019		NCGS		
2020		NATSISS	NATSISS	
2021			NCGS	
2022			NCGS	
2023			AATSIHS	AATSIHS
...				...

Source: Biddle and Yap (2010a).

In essence, Cohort 1 would be given the NATSISS questionnaire in 2014, the NCGS in 2015 and 2016, and the AATSIHS questionnaire in 2017. Cohort 2 would be given the AATSIHS questionnaire in 2017, the NCGS in 2018 and 2019 and the NATSISS questionnaire in 2020. National estimates for the 2014 NATSISS would use Cohort 1, while national estimates from the 2017 AATSIHS would use Cohort 2. This would then be repeated using Cohort 3 and Cohort 4.

The benefits of the above structure are threefold. Firstly, it would be possible for the first time to undertake robust longitudinal analysis of a core set of Indigenous outcomes. Of course, the analysis would be restricted to the core set of questions that are available on the NATSISS, the AATSIHS and the new NCGS. Major aspects of the Closing the Gap agenda, such as employment, education and health, as well as some of their determinants, such as housing, crime and mobility, would therefore be priority data items. Indigenous-specific measures of wellbeing should also, ideally, be included. The second benefit of the above structure (as opposed to a single longitudinal study) would be that the initial sample for the major surveys would still be nationally representative. The third major benefit is that, by overlapping the cohorts, the representativeness of the longitudinal aspects of the cohorts could be tested against the nationally representative survey and the outcomes in one of the surveys could be compared to predictors from the other survey.

The Closing the Gap agenda on life expectancy is generational. A part of the evidence base is analysis of existing data sets. However, this alone will not be enough. Proper randomised controlled trials that test for the causal effect of specific policies is a part of this process (as discussed below). So too is a longitudinal survey specifically designed for the Indigenous population that contains data across the life course. While the specific details will be open to debate, the above suggestion is a potential framework for a relatively cost-effective modification to the current statistical environment that would provide a yearly track of progress in meeting the Closing the Gap targets and would also allow for the development of a more robust evidence base to support Indigenous policy in Australia.



Enhanced publicly available evaluations

In her paper at the Productivity Commission's Roundtable on Better Indigenous Policies, Professor Deborah Cobb-Clark asked whether the paucity of Indigenous-specific evaluations has limited the policy and research communities' ability to learn from previous and existing policies and programs. Her response is that the 'continuing gap in Indigenous versus non-Indigenous outcomes in the face of the very substantial resources committed to Indigenous policy clearly indicates that we must do better at finding effective policies that will truly improve the wellbeing of Indigenous Australians' and that 'Program evaluation that is well done, methodologically sound, and corresponds to accepted scientific principles is critical to achieving that goal.' (Cobb-Clark 2013:86).

A consistent conclusion from papers written for the Closing the Gap Clearinghouse is that we currently do not have the data to understand what works in Indigenous policy and what doesn't work. We can learn some things from administrative data on service usage and some things from longitudinal databases. We can also gain insight from evaluations carried out on the rest of the Australian population (though there is also a gap in this area) or evaluations carried out overseas. But given the specifics of Indigenous history, culture and circumstance, there is no substitute for ethical and robust trials that properly take into account the selection into particular programs.

Support for Indigenous quantitative researchers

The final recommendation in this paper relates less to the uses of statistics about Indigenous people or the data sets that are currently available and more to those individuals and groups who analyse the data. Maggie Walter and Chris Andersen are 2 Indigenous academics at the University of Tasmania and the University of Alberta respectively. They argue in their recent book *Indigenous statistics: a quantitative research methodology* that 'Quantitative research appears to have a significant image problem within Indigenous research circles' and that 'Not only are there very few Indigenous practitioners of quantitative research, but ... for many researchers the concept of quantitative methodology itself is viewed as anathema to appropriate Indigenous research practice' (Walter & Andersen 2013:130). They argue that the issue is less to do with methods (or the underlying maths and equations) and more to do with methodology (or the assumptions and approaches driving statistical research).

Walter and Andersen's (2013) audience is mainly individual Indigenous and non-Indigenous researchers, and the target of their argument is the approach researchers take to teaching and using statistical methods on Indigenous people. However, there is also a policy implication of their argument that deserves to be emphasised. Namely, that there is a role for government and academic institutions to assist and support Indigenous quantitative researchers at the start of their career through training and scholarships, at the middle of their career through fellowships and support, and as their career progresses through access of these researchers to the policy process.

Support for quantitative researchers could be done in part through a nationally co-ordinated training program that provides support to Indigenous data collectors, data managers and analysts. Much of the analysis on Indigenous Australians using quantitative data will, however, continue to be done by non-Indigenous researchers and government employees. Spare capacity in a nationally co-ordinated training program should therefore be made available for non-Indigenous individuals working in the field, and non-Indigenous quantitative researchers who work on Indigenous issues would benefit substantially from a thorough reading of and engagement with the issues discussed in Walter & Andersen (2013).

Both approaches together will not only improve the realism and relevance of data about and for Indigenous Australians, but also the policy process. However, training programs alone aren't sufficient to maximise the likelihood that available data are used for the maximum benefit of Indigenous Australians nor that the data that are collected are sufficient for this aim. More structured processes like the National Advisory Group on Aboriginal and Torres Strait Islander Health Information and Data are needed; but they need to be extended to other domains. Perhaps more importantly, there is a need for government to commit to using and supporting robust quantitative and qualitative data.



References

- ABS (Australian Bureau of Statistics) 2000. Report on experimental Indigenous socioeconomic disadvantage indexes. Consultant's report for the CGC. In: CGC (Commonwealth Grants Commission) 2001. Report on Indigenous funding. Vol. 3 Consultants' reports. Canberra: CGC. Viewed 17 March 2014, <https://www.cgc.gov.au/index.php?option=com_content&view=article&id=53:2001-indigenous-funding-inquiry&catid=39&Itemid=160>.
- ABS 2009. Experimental estimates and projections, Aboriginal and Torres Strait Islander Australians, 1991 to 2021. ABS cat. no. 3238.0. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/4EF9B192CB67360CCA256F1B0082C453>>.
- ABS 2011. Census dictionary, 2011. ABS cat. no. 2901.0. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/2901.0>>.
- ABS 2012. Census of Population and Housing – details of undercount, 2011. ABS cat. no. 2940.0. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/2940.0>>.
- ABS 2013a. Estimates of Aboriginal and Torres Strait Islander Australians, June 2011. ABS cat. no. 3238.0.55.001. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/3238.0.55.001>>.
- ABS 2013b. Government financial estimates, Australia, 2013–14. ABS cat. no. 5501.0.55.001. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/5501.0.55.001>>.
- ABS 2013c. Introduction. In: ABS. Information paper: Australian Census longitudinal dataset, methodology and quality assessment, 2006–2011. ABS cat. no. 2080.5. Canberra: ABS. Viewed 17 March 2014, <<http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2080.5Chapter12006-2011>>.
- AHMAC (Australian Health Ministers' Advisory Council) 2012. Aboriginal and Torres Strait Islander Health Performance Framework 2012 Report. Canberra: AHMAC. Viewed 27 May 2014, <[http://www.health.gov.au/internet/main/Publishing.nsf/Content/F766FC3D8A697685CA257BF0001C96E8/\\$File/hpf-2012.pdf](http://www.health.gov.au/internet/main/Publishing.nsf/Content/F766FC3D8A697685CA257BF0001C96E8/$File/hpf-2012.pdf)>.
- AIATSIS (Australian Institute of Aboriginal and Torres Strait Islander Statistics) & FATSIL (Federation of Aboriginal and Torres Strait Islander Languages) 2005. National Indigenous Languages Survey report 2005. Canberra: Australian Government Department of Communications, Information Technology and the Arts. Viewed 27 April 2014, <<http://arts.gov.au/sites/default/files/pdfs/nils-report-2005.pdf>>.
- AIFS (Australian Institute of Family Studies) 2011. The Longitudinal Study of Australian Children data user guide, August 2011. Melbourne: AIFS. Viewed 17 March 2014, <<http://www.growingupinaustralia.gov.au/data/docs/userguide.pdf>>.
- AIHW (Australian Institute of Health and Welfare) 2010. National best practice guidelines for collecting Indigenous status in health data sets. Cat. no. IHW 29. Canberra: AIHW. Viewed 28 May 2014, <<http://www.aihw.gov.au/publication-detail/?id=6442468342>>.
- AIHW 2012. Aboriginal and Torres Strait Islander identification in community services data collections: an updated data quality report. Cat. no. IHW 80. Canberra: AIHW. Viewed 17 March 2014, <<http://www.aihw.gov.au/publication-detail/?id=60129542148>>.
- AIHW 2013a. Aboriginal and Torres Strait Islander health services report 2011–12: Online Services Report—key results. Cat. no. 104. Canberra: AIHW. Viewed 28 May 2014, <<http://www.aihw.gov.au/publication-detail/?id=60129544671>>.
- AIHW 2013b. Healthy for Life. Aboriginal Community Controlled Health Services: report card. Cat. no. IHW 97. Canberra: AIHW. Viewed 28 May 2014, <<http://www.aihw.gov.au/publication-detail/?id=60129543590>>.



AIHW 2014. National key performance indicators for Aboriginal and Torres Strait Islander primary health care: first national results June 2012 to June 2013. National key performance indicators for Aboriginal and Torres Strait Islander primary health care series. Cat. no. IHW 123. Canberra: AIHW. Viewed 28 May 2014, <<http://www.aihw.gov.au/publication-detail/?id=60129546941>>.

Altman JC & Liu J 1994. Socioeconomic status at the ATSI regional level, 1986 and 1991: data for regional planning? CAEPR discussion paper no. 76/1994. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/DP/1994DP76.php>>.

Altman JC & Nieuwenhuysen J 1979. The economic status of Australian Aborigines. Cambridge: Cambridge University Press.

Altman JC & Taylor J (eds) 1996. The 1994 National Aboriginal and Torres Strait Islander Survey: findings and future prospects. CAEPR research monograph no. 31. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 27 April 2014, <<http://caepr.anu.edu.au/Publications/mono/1996RM11.php>>.

Auspoll 2013. Australian Reconciliation Barometer 2012. Prepared by Auspoll Pty Ltd for Reconciliation Australia. Canberra: Reconciliation Australia. Viewed 21 March 2014, <<http://www.reconciliation.org.au/wp-content/uploads/2013/12/2012-Australian-Reconciliation-Barometer-Report-by-Auspoll.pdf>>.

Australian Government 2010. Closing The Gap: Prime Minister's report 2010. Canberra: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs. Viewed 27 April 2014, <<http://www.dss.gov.au/our-responsibilities/indigenous-australians/publications-articles/closing-the-gap/closing-the-gap-prime-ministers-report-2010>>.

Bailie R, Griffin J, Laycock A, Kelaher M, McNear T, Percival N et al. 2013. Sentinel Sites Evaluation: a place-based evaluation of the Indigenous Chronic Disease Package 2010–2012. Summary report. Report prepared by Menzies School of Health Research. Canberra: Australian Government Department of Health and Ageing. Viewed 27 April 2014, <<http://www.naccho.org.au/download/aboriginal-health/Sentinel%20Sites%20Evaluation%20Summary%20Report.pdf>>.

Banks G 2013. Introduction—why this roundtable? In: Productivity Commission. Better Indigenous policies: the role of evaluation. Roundtable proceedings. Canberra: Productivity Commission, pp 17–19. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies>>.

Biddle N 2009a. Ranking regions: revisiting an index of relative Indigenous socioeconomic outcomes. CAEPR working paper no. 50/2009. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/WP/2009WP50.php>>.

Biddle N 2009b. Same suburb, different worlds: comparing Indigenous and non-Indigenous outcomes in city suburbs and large regional towns. *People and Place* 17(2):56–63. Viewed 19 March 2014, <<http://arrow.monash.edu.au/vital/access/manager/Repository/monash:64174>>.

Biddle N 2013. Socioeconomic outcomes. CAEPR Indigenous Population Project: 2011 Census paper no. 13/2013. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/census-papers/2013CP13.php>>.

Biddle N & Wilson T 2013. Indigenous Australian population projections: problems and prospects. *Journal of Population Research* 30(2):101–16. doi:10.1007/s12546-013-9104-2.

Biddle N & Yap M 2010a. Closing the (data) Gap. *Indigenous Law Bulletin* 7(19). Viewed 19 March 2014, <<http://www.austlii.edu.au/au/journals/ILB/2010/26.html>>.

Biddle N & Yap M 2010b. Demographic and socioeconomic outcomes across the Indigenous Australian lifecycle: evidence from the 2006 Census. CAEPR research monograph no. 31. Canberra: ANU E Press. Viewed 19 March 2014, <http://epress.anu.edu.au/c31_citation.html>.



- Blair EM, Zubrick SR & Cox AH 2005. The Western Australian Aboriginal Child Health Survey: findings to date on adolescents. *Medical Journal of Australia* 183(8):433–5. Viewed 19 March 2014, <<https://www.mja.com.au/journal/2005/183/8/western-australian-aboriginal-child-health-survey-findings-date-adolescents>>.
- Blinder AS 1973. Wage discrimination: reduced form and structural estimates. *Journal of Human Resources* 8(4):436–55. doi:10.2307/144855.
- Booth AL & Carroll N 2008. Economic status and the Indigenous/non-Indigenous health gap. *Economics Letters* 99(3):604–6. doi:10.1016/j.econlet.2007.10.005.
- Campbell DT 1969. Reforms as experiments. *American Psychologist* 24(4):404–29. doi:10.1037/h0027982.
- CGC (Commonwealth Grants Commission) 2010. Report on GST revenue sharing relativities—2010 review. Volume 1: Main report. Canberra: CGC. Viewed 19 March 2014, <https://cgc.gov.au/attachments/article/27/2010_Review_final_report_vol_1.pdf>.
- Cobb-Clark DA 2013. Chapter 5: The case for making public policy evaluations public. In: Productivity Commission. *Better Indigenous policies: the role of evaluation. Roundtable proceedings*. Canberra: Productivity Commission, pp 81–91. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies/>>.
- CRAETP (Committee of Review of Aboriginal Employment and Training Programs) 1985. Report of the Committee of Review of Aboriginal Employment and Training Programs. Canberra: Australian Government Publishing Service. Viewed 27 April 2014, <<http://www.voced.edu.au/content/ngv29210>>.
- CtGC (Closing the Gap Clearinghouse) 2013. What works to overcome Indigenous disadvantage: key learnings and gaps in the evidence 2011–12. Produced for the Closing the Gap Clearinghouse. AIHW cat. no. IHW 107. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies. Viewed 19 March 2014, <<http://www.aihw.gov.au/closingthegap/publications/>>.
- FaHCSIA (Australian Government Department of Families, Housing, Community Services and Indigenous Affairs) 2009. Footprints in time: the Longitudinal Study of Indigenous Children—key summary report from Wave 1. Canberra: FaHCSIA. Viewed 19 March 2014, <<http://www.dss.gov.au/our-responsibilities/indigenous-australians/publications-articles/families-children/footprints-in-time-the-longitudinal-study-of-indigenous-children-Isic/key-summary-report-from-wave-1-2009>>.
- FaHCSIA 2013. Footprints in time: the Longitudinal Study of Indigenous Children—report from Wave 4. Canberra: FaHCSIA. Viewed 20 March 2014, <<http://www.dss.gov.au/our-responsibilities/indigenous-australians/publications-articles/families-children/footprints-in-time-the-longitudinal-study-of-indigenous-children-Isic/key-summary-report-from-wave-4>>.
- Gertler PJ, Martinez S, Premand P, Rawlings LB & Vermeersch CMJ 2010. *Impact evaluation in practice*. Washington, DC: The International Bank for Reconstruction and Development/The World Bank. Viewed 20 March 2014, <http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf>.
- Gray MC & Auld AJ 2000. Towards an index of relative Indigenous socioeconomic disadvantage. CAEPR discussion paper no. 196/2000. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/DP/2000DP196.php>>.
- Harrison LJ, Goldfeld S, Metcalfe E & Moore T 2012. Early learning programs that promote children’s developmental and educational outcomes. Resource sheet no. 15. Produced for the Closing the Gap Clearinghouse. AIHW cat. no. IHW 76. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies. Viewed 19 March 2014, <<http://www.aihw.gov.au/closingthegap/publications/>>.



Helme S & Lamb S 2011. Closing the school completion gap for Indigenous students. Resource sheet no. 6. Produced for the Closing the Gap Clearinghouse. AIHW cat. no. IHW 38. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies. Viewed 19 March 2014, <<http://www.aihw.gov.au/closingthegap/publications/>>.

Hewitt B 2012. The Longitudinal Study of Indigenous Children: implications of the study design for analysis and results. Prepared for the Department of Families, Housing, Community Services and Indigenous Affairs. Brisbane: Institute for Social Science Research, The University of Queensland. Viewed 21 March 2014, <<http://www.dss.gov.au/sites/default/files/files/about-fahcsia/publication-articles/footprints/Implications%20of%20the%20study%20design%20for%20analysis%20and%20results.pdf>>.

Hunter B & Biddle N (eds) 2012. Survey analysis for Indigenous policy in Australia: social science perspectives. CAEPR research monograph no. 32. Canberra: ANU E Press. Viewed 19 March 2014, <<http://press.anu.edu.au/titles/centre-for-aboriginal-economic-policy-research-caepr/survey-analysis-for-indigenous-policy-in-australia/pdf-download/>>.

Hunter BH (ed.) 2006. Assessing the evidence on Indigenous socioeconomic outcomes: a focus on the 2002 NATSISS. CAEPR research monograph no. 26. Canberra: ANU E Press. Viewed 27 April 2014, <http://press.anu.edu.au/titles/centre-for-aboriginal-economic-policy-research-caepr/c26_citation/pdf-download/>.

James M 2013. Chapter 7: Designing evaluation strategies. In: Productivity Commission. Better Indigenous policies: the role of evaluation. Roundtable proceedings. Canberra: Productivity Commission, pp 107–18. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies>>.

Jordan K, Bulloch H & Buchanan G 2010. Statistical equality and cultural difference in Indigenous wellbeing frameworks: a new expression of an enduring debate. *Australian Journal of Social Issues* 45(3):333–62. Viewed 20 March 2014, <<http://search.informit.com.au/documentSummary;dn=414770921821552;res=IELFSC>>.

Kalb G, Le T, Hunter B & Leung F 2012. Decomposing differences in labour force status between Indigenous and non-Indigenous Australians. IZA discussion paper no. 6808. Bonn: Institute for the Study of Labor (IZA). Viewed 27 April 2014, <<http://ftp.iza.org/dp6808.pdf>>.

Leigh A 2010. Chapter 10: Evidence-based policy: summon the randomistas? In: Productivity Commission. Strengthening evidence-based policy in the Australian Federation. Volume 1: Proceedings. Roundtable proceedings. Canberra: Productivity Commission, pp 215–26. Viewed 20 March 2014, <<http://www.pc.gov.au/research/conference-proceedings/strengthening-evidence>>.

Malezer L 2013. Chapter 4: Challenges in evaluating Indigenous policy. In: Productivity Commission. Better Indigenous policies: the role of evaluation. Roundtable proceedings. Canberra: Productivity Commission, pp 69–79. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies>>.

Martin K (Mirraboopa B) 2003. Ways of knowing, being and doing: a theoretical framework and methods for indigenous and indigenist research. *Journal of Australian Studies*, 27(76), 203–214. Viewed 3 June 2014 <<http://eprints.qut.edu.au/7182/1/7182.pdf>>.

NHMRC (National Health and Medical Research Council) 2000. How to use the evidence: assessment and application of scientific evidence. Handbook series on preparing clinical practice guidelines. Canberra: NHMRC. Viewed 19 March 2014, <<https://www.nhmrc.gov.au/guidelines/publications/cp69>>.

Oaxaca R 1973. Male–female wage differentials in urban labour markets. *International Economic Review* 14(3):693–709. doi:10.2307/2525981.

PC (Productivity Commission) 2013. Better Indigenous policies: the role of evaluation. Roundtable proceedings. Canberra: PC. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies>>.



Rao V & Woolcock M 2003. Chapter 8: Integrating qualitative and quantitative approaches in program evaluation. In: Bourguignon F & Pereira da Silva LA (eds). *The impact of economic policies on poverty and income distribution: evaluation techniques and tools*. New York: Oxford University Press, pp 165–90. Viewed 20 March 2014, <http://siteresources.worldbank.org/INTPSIA/Resources/490023-1121114603600/12930_Chapter8.pdf>.

Ravallion M 2008. Evaluating anti-poverty programs. In: Schultz TP & Strauss J (eds). *Handbook of development economics*. Vol. 4. Amsterdam: North-Holland. Viewed 20 March 2014, <http://siteresources.worldbank.org/INTISPMA/Resources/383704-1153333441931/Evaluating_Antipoverty_Programs.pdf>.

RCAP (Royal Commission on Aboriginal Peoples) 1996. *Report of the Royal Commission on Aboriginal Peoples. Volume 5: Renewal: a twenty-year commitment*. Ottawa: Canada Communication Group—Publishing. Viewed 27 April 2014, <http://caid.ca/Vol_5_RepRoyCommAborigPple.html>.

Ross R 1993. A probit analysis of Aboriginal employment prospects in New South Wales. *Economic Record* 69(3):253–8. doi:10.1111/j.1475-4932.1993.tb02105.x.

Rowley CD 1970. *The destruction of Aboriginal society. Aboriginal policy and practice—Volume I*. Canberra: Australian National University Press.

Rowley CD 1971a. *Outcasts in white Australia. Aboriginal policy and practice—Volume II*. Canberra: Australian National University Press.

Rowley CD 1971b. *The remote Aborigines. Aboriginal policy and practice—Volume III*. Canberra: Australian National University Press.

Rowse T 2012. *Rethinking social justice: from ‘peoples’ to ‘populations’*. Canberra: Aboriginal Studies Press.

SCRGSP (Steering Committee for the Review of Government Service Provision) 2011. *Overcoming Indigenous disadvantage: key indicators 2011*. Canberra: Productivity Commission. Viewed 20 March 2014, <<http://www.pc.gov.au/gsp/overcoming-indigenous-disadvantage/key-indicators-2011>>.

SCRGSP 2012. *2012 Indigenous expenditure report*. Canberra: Productivity Commission. Viewed 20 March 2014, <<http://www.pc.gov.au/gsp/ier/indigenous-expenditure-2012/report>>.

SCRGSP 2013. *Report on government services 2013: Indigenous compendium*. Canberra: Productivity Commission. Viewed 27 April 2014, <<http://www.pc.gov.au/gsp/rogs/indigenous-compendium-2013>>.

Smith LT 1999. *Decolonizing methodologies: research and Indigenous peoples*. London and New York: Zed books Ltd.

Taylor J 2013. Chapter 8: Data for better Indigenous policy evaluation: achievements, constraints and opportunities. In: Productivity Commission. *Better Indigenous policies: the role of evaluation. Roundtable proceedings*. Canberra: Productivity Commission, pp 119–30. Viewed 17 March 2024, <<http://www.pc.gov.au/research/conference-proceedings/better-indigenous-policies>>.

Taylor J & Stanley O 2005. *The opportunity costs of the status quo in the Thamarrurr region*. CAEPR working paper no. 28/2005. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/WP/2005WP28.php>>.

Taylor J, Doran B, Parriman M & Yu E 2012. *Statistics for community governance: the Yawuru Indigenous Population Survey of Broome*. CAEPR working paper no. 82/2012. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/WP/2012WP82.php>>.



Tesfaghiorghis H 1991. Geographic variations in the economic status of Aboriginal people: a preliminary investigation. CAEPR discussion paper no. 2/1991. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 19 March 2014, <<http://caepr.anu.edu.au/Publications/DP/1991DP2.php>>.

Walter M & Andersen C 2013. Indigenous statistics: a quantitative research methodology. Walnut Creek, California: Left Coast Press.

Yap M & Biddle N 2010. Gender gaps in Indigenous socioeconomic outcomes: Australian regional comparisons and international possibilities. *International Indigenous Policy Journal* 1(2). Viewed 20 March 2014, <<http://ir.lib.uwo.ca/iipj/vol1/iss2/3>>.

Yu P 2011. The power of data in Aboriginal hands. Paper presented to Social Science Perspectives on the 2008 National Aboriginal and Torres Strait Islander Social Survey conference, The Australian National University, 11–12 March. Canberra: Centre for Aboriginal Economic Policy Research, The Australian National University. Viewed 21 March 2014, <http://caepr.anu.edu.au/Seminars/11/NATSIS-2011/11_4_Seminar.php>.

Acknowledgments

Dr Nicholas Biddle is a Fellow at the Centre for Aboriginal Economic Policy Research, Australian National University. He is grateful to Clearinghouse staff and officers from relevant government departments who provided comments on an earlier version of this paper. Dr Biddle would also like to thank (in chronological order) Kaely Woods, Katherine Curchin, Boyd Hunter, Francis Morphy, Maggie Brady, Jon Altman, Heather Crawford, Talia Avrahamzon and Mandy Whitford for their individual comments.

Abbreviations

AATSIHS	Australian Aboriginal and Torres Strait Islander Health Survey
ABS	Australian Bureau of Statistics
ACLD	Australian Census Longitudinal Dataset
AEDI	Australian Early Development Index
CGC	Commonwealth Grants Commission
COAG	Council of Australian Governments
HILDA	Household, Income and Labour Dynamics Australia
HPF	Health Performance Framework
ICDP	Indigenous Chronic Disease Package
LSAC	Longitudinal Study of Aboriginal Children
LSAY	Longitudinal Study of Aboriginal Youth
LSIC	Longitudinal Study of Indigenous Children
NATSIS	National Aboriginal and Torres Strait Islander Survey
NATSISS	National Aboriginal and Torres Strait Islander Social Survey
NCGS	National Closing the Gap Survey
OID	Overcoming Indigenous Disadvantage



Terminology

Indigenous: 'Aboriginal and Torres Strait Islander' and 'Indigenous' are used interchangeably to refer to Australian Aboriginal and Torres Strait Islander people. The Closing the Gap Clearinghouse uses the term 'Indigenous Australians' to refer to Australia's first people. This term includes 'Aboriginal Australians' and 'Torres Strait Islander people'.

Funding

This paper was commissioned by the Closing the Gap Clearinghouse. The Clearinghouse is a Council of Australian Governments' initiative jointly funded by all Australian Governments. The Australian Institute of Health and Welfare in collaboration with the Australian Institute of Family Studies deliver the Clearinghouse.

Suggested citation

Biddle N 2014. Data about and for Aboriginal and Torres Strait Islander Australians. Issues paper no. 10. Produced for the Closing the Gap Clearinghouse. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies.

Copyright

© Australian Institute of Health and Welfare 2014 

This product, excluding the AIHW logo, Commonwealth Coat of Arms and any material owned by a third party or protected by a trademark, has been released under a Creative Commons BY 3.0 (CC BY 3.0) licence. Excluded material owned by third parties may include, for example, design and layout, images obtained under licence from third parties and signatures. We have made all reasonable efforts to identify and label material owned by third parties.

You may distribute, remix and build upon this work. However, you must attribute the AIHW as the copyright holder of the work in compliance with our attribution policy available at <www.aihw.gov.au/copyright/>. The full terms and conditions of this licence are available at <<http://creativecommons.org/licenses/by/3.0/au/>>.

Enquiries relating to copyright should be addressed to the Head of the Media and Strategic Engagement Unit, Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601.

ISBN 978-1-74249-601-6

ISSN 2201-845X

Cat. no. IHW 136