ABS Submission to the Productivity Commission's Inquiry into a National Education Evidence Base

Summary

This submission addresses a range of issues that should be a focus of attention when endeavouring to enhance the national education evidence base. In particular:

1. Increasing the utilisation of existing data to inform complex policy issues, particularly through integrating existing administrative and statistical datasets
2. Leveraging opportunities to improve consistency and co-ordination of statistical methodology and approaches deployed across the education sector; and
3. The need to strengthen co-operative arrangements to address national statistical needs in the context of challenges imposed by a sector featuring many stakeholders and their diverse interests.

For inquiries about this submission please contact Michelle Marquardt, Program Manager, Education, Crime and Culture Branch

Introduction

The Australian Bureau of Statistics (ABS) welcomes the opportunity to contribute to the Productivity Commission's Inquiry and is supportive of the aim to improve information for policy and research purposes through integration and increased utilisation of the various rich datasets available in the education sector.

The ABS is Australia's official national statistical agency. The role of the ABS is to assist and encourage informed decision-making, research and discussion within governments and the community by leading a high quality, objective and responsive national statistical service.

Identifying and responding to the information needs of governments and the community is a high priority for the ABS. The ABS provides information in a number of ways including data freely available on our website, customised datasets, as well as a range of products which allow users to explore in greater depth the wealth of information ABS has collected.

The ABS produces a range of education statistics including from the *Australian Census of Population and Housing*, surveys and administrative data sources. The routine collection of education indicators in all ABS household collections enables a broad range of analysis across a wide range of indicators including how education interrelates with labour force status, health, disability, income and housing. The ABS believes that more could be done through the integration of data sources, particularly for

the purpose of longitudinal outcomes analysis and more in-depth and robust analyses not possible from the datasets on their own.

The concept of a National Education Information Agreement is supported by the ABS, noting such an agreement has been proposed in the submission to this inquiry made by the Australian Government Department of Education and Training. It is expected that such an agreement would act to clarify roles and responsibilities amongst stakeholders and prioritise national statistical outputs which would better focus and leverage the work needed to improve the evidence base.

Scope of the Inquiry

When considering the overall scope of the Inquiry, the ABS sees value in extending the scope beyond schools and school education. Integration of datasets relating to before and after school experiences will enable more robust analysis of outcomes achieved through early childhood and school education.

Objectives and Framework

The main objective of a National Education Evidence Base would be to improve educational outcomes through improving the evidence base for policy makers and researchers. Including data sources, such as the *Australian Census of Population and Housing*, can enable the creation of rich sources that could provide a holistic view of the educational journey. Broadening the evidence base to consider other aspects of the life cycle, such as labour force, personal and family characteristics, and other socio-economic factors which interrelate with educational outcomes will support more rounded analyses to be undertaken.

Associated Costs and Benefits

At present, there exist a variety of rich sources of administrative data that could be integrated[1] with each other and with other sources such as the Census to deliver benefits beyond those that could be obtained through the investigation of each dataset individually.

The ABS believes there are many potential benefits arising from the integration of datasets, particularly at the national level. These include:

- Consistent treatment of data through the integration process ensuring robust, comparable output. This in turn means that opportunities to compare natural experiments (for example exploring the impact of differences in school starting age) can be done efficiently and effectively;
- Consistent metadata;
- Streamlined release and access practices;

[1] Data integration means bringing information together. It is an efficient and effective way of creating new insights from reusing existing data to address important questions about Australian society. When a problem is identified that no single set of data can resolve, data can be brought together in a very safe and controlled way, to enable analysis of the issue.

- Establishment of a 'one stop shop' model, where large volumes of subject matter are held in the one location supporting effective knowledge of, and easier access to, data and statistics.

In the ABS' experience, the costs and barriers to data integration include:

- Enabling data access due to real and perceived legislative barriers (in terms of gaining access for linkage and in terms of providing access to microdata for researcher and policy analysts, and releasing aggregate results and analysis of the linked information);

- Adequate funding to support the management, acquisition, transformation, linkage and dissemination of data.

The ABS notes recent initiatives to test and enable access and data integration of education data from various sources and the Public Sector Data Management project of the Department of Prime Minister and Cabinet.

Data sharing

The main challenge to implement data linkage in the education sector are the legislation and policy barriers, real or perceived, associated with data sharing. Such barriers can significantly increase the complexity and costs associated with developing and implementing data integration projects. In addressing these barriers, the ABS has had some success in recent years in partnering with other departments and agencies with similar interests in growing the statistical evidence base.

Positive examples of these partnerships include:

- Working with the Australian, Queensland and Tasmanian Government departments to integrate and analyse data from the 2011 *Australian Census of Population and Housing*, the *National Early Childhood Education and Care Collection* (NECECC), the *Australian Early Development Census (AEDC), National Assessment Program – Literacy and Numeracy* (NAPLAN), and school enrolments. This resulted in a series of articles published by the ABS with new insights about the factors influencing childhood development and student achievement (see catalogues 4261.3 and 4261.6).

- The integration and analysis of NECECC, AEDC and National Quality Assessment data sets for the Australian and State and Territory Governments. These data sets are being used to identify critical factors in the successful transition of children between preschool and the first year of schooling. The project has been undertaken to assist senior officials with implementation of the National Aboriginal and Torres Strait Islander Education Strategy.

Access to ABS integrated data sets like those mentioned above has also been significantly improved in the past year with the development of the ABS' DataLab, which is a secure and safe (remote and onsite) computer environment in which researchers can analyse de-identified data using their choice of software package (e.g. SAS, SPSS, etc.). Partnerships have at times included utilising 'inpostings' whereby an ABS officer works within another organisation to facilitate understanding and communication particularly with respect to tricky data access and analysis issues, including through the DataLab products.

Privacy

The ABS acknowledges that concerns about privacy can act as a significant barrier that inhibits data sharing and statistical development in the education sphere. Whilst concerns relating to compliance with relevant privacy legislation and the protection of rich and extremely sensitive data pertaining to children (and the barriers they create) are real and significant, the ABS believes that they can be overcome.

Key education data sources, particularly those with state and territory custodians, are governed by jurisdiction specific privacy related legislation. Similarly national agencies that may collect or take custodianship of data for statistical purposes have responsibilities under law (*Privacy Act, 1988* and the Australian Privacy Principles) and their own legislation (in the ABS case, the *Census and Statistics Act 1905*). However, it is the ABS' general view that most prevailing legislation is able to support safe collection in regards to privacy, compilation and transmission of data and development and dissemination of appropriate statistics. In the ABS' case the *Census and Statistics Act 1905* ensures that the collection, processing and release of data is all undertaken in a privacy-conscious and safe way.

As statistical developments such as data integration featuring the linkage of multi-jurisdictional data proceeds, it is inevitable and appropriate that the management of this rich data is scrutinised. It should be acknowledged that there are issues that need to be considered. Examples include challenges to the ability to offer informed consent to those (children and families) whose data is repurposed for statistical use, similarly access to richer and more detailed data can offer more inroads and incentives towards data re-identification. Again, however, the controls available to and imposed by key agencies involved with this work would seem adequate and commensurate with the privacy risks. The health domain may offer examples of enabling such initiatives.

The ABS has reviewed its policies and practices and is taking a more holistic approach to data release that includes consideration of the 5 safes principles model[2]. It has also been trialling improved access, allowing for longer pre-release periods, and the introduction of a trial Virtual ABS DataLab. The *Preschool Education, Australia, 2015* microdata has recently been made available through this system, receiving positive feedback from key clients.

Data comparability

There are a number of comparability issues in the national education and training data collections, pertaining mostly to the number of disparate systems used for data collection, variance in service starting ages (varying across jurisdictions and across sectors), and differences in scope for particular data collections. Reference periods also differ across the collections. These issues are not insignificant, where the national coherence of data is desired. These differences contribute to making COAG National Partnership Agreement and other performance reporting a challenge where

[2] For explanation see:

- ABS Information Paper: Transforming Statistics for the Future, Feb 2016 (cat. no. 1015.0), http://www.abs.gov.au/AUSSTATS/abs@.nsf/be4aa82cd8cf7f07ca2570d60018da27/e4d483bab4e1ad93ca257f4c00170bb6!OpenDocument,
- Desai, T., Ritchie, F., Welpton, R. (2016) *Five Safes: designing data access for research;* University of the West of England, Economics Working Paper Series 1601; http://rsss.anu.edu.au/sites/default/files/Ritchie_5safes.pdf ,
- Ritchie, F. (2013) *International access to restricted data: A principles-based standards approach*; Statistical Journal of the IAOS 29 (2013) 289-300; http://content.iospress.com/download/statistical-journal-of-the-iaos/sji00780?id=statistical-journal-of-the-iaos%2Fsji00780

understanding the broader sector, jurisdiction, regional or sub-population context can be very important in understanding the drivers of difference in headline performance indicators. On the other hand, analysing the impact of such differences can lead to valuable educational and other social research outcomes.

Technology

The ABS strongly agrees with the proposition that technological advancement is a critical enabler to enhancing the quality and range of the evidence base. In recognition of this importance, recently the ABS sought, and was successful in obtaining significant funding from the Australian Government to replace aging infrastructure with new systems and technologies to support improvements across all stages of the process for developing data and statistics. This funding has helped drive new technological developments as well as supporting existing technological innovation that offer multiple advantages to both the ABS and to other stakeholders (data custodians and users) involved in the production of statistics. Some examples of breakthroughs recently emerging that are particularly relevant to the education field and have featured technology as a key enabler include:

- The capacity to expand the potential inherent in existing raw data sources when turning them into statistics and information. Particularly when integrating large datasets from different sources, methodological, technical and processing power advances with both software and hardware have enabled better integration delivering high quality and rich datasets for analysis.
- The capacity to deliver data more directly to the analyst or research user, in greater detail than previously available, while continuing to safeguard the privacy of the entities to which the data pertains. Microdata access products such as ABS TableBuilder supporting the delivery of detailed tabular and graphical outputs have significantly improved the depth and breadth of specific outputs from individual statistical collections. Similarly the ABS is currently pioneering the delivery of de-identified unit level data to the user's desktop using a secure virtual application.
- Developing electronic instruments for the collection of survey and administrative data at source which can dramatically reduce costs of collection, compilation and editing.
- Developing automated methods for confidentialising data to protect confidentiality of individual's information.

There are, however, some significant barriers associated with the introduction of technology. In particular, it should be acknowledged that advancing technology alone offers limited returns without addressing the other barriers that hinder data and statistical development that are mentioned in this submission, even where the required data flows are relatively easy to manage from a technical perspective. Even within the ABS, technological enablement has needed policy support by resetting certain internal policy settings before the organisation could leverage all the advantages that innovation offered.

A key feature of the Australian Education sector is that it features many stakeholders with significant shared responsibility for delivery of data. Data custodians have different systems for collecting and compiling data, which can be subject to (for example) different policies and legislation in different states and territories (particularly legislation related to privacy and the resulting implications when repurposing of administrative data for statistical production). Investment is needed in areas such as setting common standards for data and in developing systems and infrastructure that will effectively "talk" with other systems. This is additional to reaching consensus both bilaterally and multi-laterally over the appropriateness of various statistical projects and statistical reporting.

Additionally, the cost and effort required to develop new technology and in such mundane tasks as digitising or transferring existing information to a new common digital format should not be underestimated.

Institutions

The ABS agrees that it is necessary that clear institutional and governance arrangements are established in the Education sector to enable the production of national statistics. This is particularly important given that the large range of institutions currently involved at both state/territory and national level with policy and/or statistical responsibility can create challenges to reach agreement over and execute statistical priorities requiring complex and coordinated action to address.

The ABS is an important contributor to national statistical arrangements noting the statistical co-ordination and service functions of the ABS as set down in the *Australian Bureau of Statistics Act 1975 S(6)*. The ABS also has significant expertise dealing with technical issues that often become a focus of national level governance such as developing effective data collection and compilation systems featuring multi sourced data, developing national statistical indicators, developing and implementing national statistical standards and handling big data.

The ABS currently, and has for many years, been the producer of key national statistics directly related to education (for example, the *National School Statistics Collection*, the *National Early Childhood Education and Care Collection* and the *Childhood Education and Care survey*) that input into national reporting frameworks. The ABS is also a producer and the custodian of other national datasets containing a wealth of information about children and youth, families, labour force, disadvantaged populations and other relevant social and economic topics. Although these datasets are covering a broader range of themes than just education, they form essential components of the education evidence base. The utilisation of these datasets offers significant potential that is starting to be demonstrated through various projects linking data over time and across the complete education life cycle which is valuable for understanding and evaluating education sector programs, policies and outcomes. Accordingly, institutional arrangements will also need to consider the utilisation of this new wave of rich and complex multi-source data.

Population and Social Statistics Division
Australian Bureau of Statistics

Selected References

Below is a selection of ABS publications, microdata products and information papers that may be of relevance to the Inquiry.

Australian Bureau of Statistics (ABS), 2016, *Information Paper: Transforming Statistics for the Future, Feb 2016, cat. no. 1015.0,* ABS, Canberra
http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1015.0Main%20Features5Feb%202016?opendocument&tabname=Summary&prodno=1015.0&issue=Feb%202016&num=&view=

Australian Bureau of Statistics (ABS), 2016, *Case Study: Education,* ABS, Canberra,
http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+Case+Study:+Education

Australian Bureau of Statistics (ABS), 2016, *ABS Data Integration,* ABS, Canberra,
http://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Statistical+Data+Integration

Australian Bureau of Statistics (ABS), 2016*, Microdata: Education and Work, May 2015, cat. no. 6227.0.30.001,* ABS, Canberra,
http://www.abs.gov.au/ausstats%5Cabs@.nsf/0/4985F229045DE899CA256ECF0001B891?Opendocument

Australian Bureau of Statistics (ABS), 2016, *Microdata: Preschool Education, Australia, 2015, cat. no. 4240.0.55.003*, ABS, Canberra,
http://www.abs.gov.au/ausstats%5Cabs@.nsf/0/D1A57857AF1EB6A3CA257CC90014404F?Opendocument

Australian Bureau of Statistics (ABS), 2016, *Schools, Australia, 2015, cat. no. 4221.0*, ABS, Canberra,
http://www.abs.gov.au/ausstats/abs@.nsf/mf/4221.0

Australian Bureau of Statistics (ABS), 2016, *Preschool Education, Australia, 2015, cat. no. 4240.0*, ABS, Canberra,
http://www.abs.gov.au/AUSSTATS%5Cabs@.nsf/0/BC7AC812771E8FEFCA2578680014F525?Opendocument

Australian Bureau of Statistics (ABS), 2015, *Microdata: Outcomes from Vocational Education and Training in Schools, Australia, 2006-2011, cat. no. 4260.0.55.001,* ABS, Canberra,
http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4260.0.55.001Explanatory%20Notes302006-2011?OpenDocument

Australian Bureau of Statistics (ABS), 2015, *Essential Statistical Assets for Australia, 2015, cat. no. 1395.0 ,* ABS, Canberra,
http://www.abs.gov.au/ausstats/abs@.nsf/0/A7FFA19FF6239724CA257F65002272E1?Opendocument

Australian Bureau of Statistics (ABS), 2015, *Childhood Education and Care, Australia, June 2014, cat. no. 4402.0 ,* ABS, Canberra,
http://www.abs.gov.au/Ausstats/abs@.nsf/0/6EF8CC70C77E0A68CA2579F30011EA6C?OpenDocument

Australian Bureau of Statistics (ABS), 2015, *Microdata: Childhood Education and Care, Australia, June 2014, cat. no. 4402.0.55.001,* ABS, Canberra,
http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/4402.0.55.001?OpenDocument

Australian Bureau of Statistics (ABS), 2015, *Education and Work, Australia, May 2015, cat. no. 6227.0*, ABS, Canberra, http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/6227.0?OpenDocument

Australian Bureau of Statistics (ABS), 2014, *Educational outcomes, experimental estimates, Tasmania, 2006-2013*, *cat. no. 4261.6*, ABS, Canberra, http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/4261.6?OpenDocument

Australian Bureau of Statistics (ABS), 2014, *Outcomes from Vocational Education and Training in Schools, Australia, 2006-2011, cat. no. 4260.0,* ABS, Canberra, http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/4260.0Main%20Features32006-2011?opendocument&tabname=Summary&prodno=4260.0&issue=2006-2011&num=&view=

Australian Bureau of Statistics (ABS), 2014, *Research Paper: Assessing the Effect of Geographic Information on Data Linkage Quality Using Vocational Education and Training in Schools Data, 2014, cat. no. 4269.0.55.001*, ABS, Canberra, http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/4269.0.55.001?OpenDocument

Australian Bureau of Statistics (ABS), 2013*, Research Paper: Assessing the Feasibility of Linking 2011 Vocational Education and Training in Schools Data to 2011 Census Data, Dec 2013*, *cat. no. 1351.0.55.044,* ABS, Canberra, http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/1351.0.55.044?OpenDocument

Australian Bureau of Statistics (ABS), 2013, *Programme for the International Assessment of Adult Competencies, Australia, 2011-12, cat. no. 4228.0,* ABS, Canberra, http://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/4228.0