

Friday, 9 December 2016

Mr Peter Harris AO  
Presiding Commissioner  
Data Availability and Use  
Productivity Commission  
GPO Box 1428  
CANBERRA ACT 2601

**Professor Margaret Harding**  
Deputy Vice-Chancellor, Research

**+61 2 6125 8487**  
[dvc.research@anu.edu.au](mailto:dvc.research@anu.edu.au)

Canberra ACT 0200 Australia  
[www.anu.edu.au](http://www.anu.edu.au)

CRICOS Provider No. 00120C

## **Re: Productivity Commission Public Inquiry into Data Availability and Use**

Dear Mr Harris

Thank you for the opportunity to comment on the Productivity Commission's draft report on Data Availability and Use. The benefits of making data available in a way that is safe and protects the confidentiality of those who provide their data to government, researchers and commercial organisations are immeasurable. Across the physical, social and life sciences, advances in knowledge are underpinned by careful data analysis and open sharing of information. Good public policy cannot occur without such knowledge.

In responding to the Productivity Commission's draft report, five specific goals and principles are outlined and discussed below:

### **Goal 1. Encouraging and supporting access to data, particularly for the research and academic community**

The draft report sets out a number of benefits to society and public policy from the research and academic community being able to access high quality data from government. An understandable focus of the draft report is on accessing administrative datasets, given the relative difficulties in accessing such data until now. The ANU is strongly supportive of making such data available.

It is also important to highlight the significant benefit of having a deeper, timely and more flexible access to survey data collected by government. This includes, but is not limited to, the ability to link survey data with other sources of information.

One aspect of data access that is covered in less detail in the Productivity Commission draft report is replication and validation studies. There is a relative lack of such studies in Australia, especially in the social and health sciences, despite the need to ensure that findings from individual studies are not one-off or driven by idiosyncrasies in the sample. This is partly driven by the difficulties of individual researchers obtaining data used by other researchers. Such access is particularly difficult for research that is carried out from within government. This will become an increasingly important issue as governments make more use of Randomised Controlled Trials in the design and delivery of programs. Any models that follow from the draft report will need to consider mechanisms to make such replication and validation studies possible.

### **Goal 2. Having a voice and representation with respect to data, including its use**

Many of the recommendations in the draft report relate to data confidentialising and data linkage. Depending on the data involved, and the level of confidentialisation required, such processes can require a high level of technical skill and expertise. Some of this expertise is held within government. However, there is also a considerable amount of expertise within the broader university and research community that will be required if the volume of data suggested by the Productivity Commission is to be made available.

There are specific aspects of the report that the ANU specifically, and the university community more broadly, would have a strong knowledge of and incentive to contribute to. One aspect that does not receive much discussion in the report is the use of unit record data for teaching and training purposes. The vast majority of researchers within government, academia and the community sector are trained at Australian universities. This training requires access to high quality data in a way that protects the confidentiality of respondents.

There is a considerable amount of further research required to ensure that the recommendations from the report are implemented in an evidence-based manner. From a social science perspective, there is a need to understand community support for data access and use, how that support varies across the community, and how that support is affected by external shocks and broader social trends. From a methodological perspective, there is an ongoing need to understand the most efficient and effective way to make data available in a way that maximises its usefulness, whilst minimising disclosure risk. Funding organisations may consider targeted calls for research into these and related issues.

More broadly, the ANU advocates the need for expert input into the finalisation and implementation of the report to ensure it is of the highest possible quality and is informed by researchers and other end-users of data. This involves supporting partnerships with policymakers, practitioners and other stakeholders with respect to data. A number of ANU staff members are currently involved in technical, and subject matter advisory committees related to data availability and use. The ANU would welcome a continuation and expansion of such a role.

### **Goal 3. Encouraging and being seen to encourage data sharing, including for data collected by ANU**

The draft report states (p 22) that *'There is a need for the research community to put its house in order when it comes to data sharing. Just as government data custodians should consider that they hold data not solely for their own purposes but in the public interest on behalf of citizens, so too should the data of publicly funded research be available beyond the initial researchers.'* The ANU strongly agrees with the statement that data from publicly funded research be made available to researchers from outside the ANU, and from outside the university sector more broadly.

There are a number of examples of largescale data collection activities undertaken by the ANU for which unit record data is made available to external researchers within a very short period of time from data collection. This includes (but is not limited to) ANU Poll, the Australian Election Study, and 45 and up. These datasets are currently made available through the Australian Data Archive (<https://www.ada.edu.au/>), a national resource that is currently funded exclusively by the ANU. Such data is used extensively by external researchers in Australian universities, within government in Australia, and internationally.

### **Goal 4. Ensuring appropriate resources to support access to data**

Draft recommendation 9.9 states that *'Public research funding should be prioritised on the basis of progress made by research institutions in making their researchers' data widely available.'* This is a recommendation that the ANU is broadly in support of, with the important caveat that funding certainty is required to support the ongoing curation and depositing of data in a way that is searchable and accessible. It is reasonably common for government funded grants (including from the ARC and NHMRC) to be granted with reduced budgets. Including separate budget items for data release and archiving, and a consistent pricing formula that takes into account the fixed costs, as well as the marginal costs associated with additional variables and observations, would be welcome by researchers and funding agencies in ensuring that appropriate consideration of data and funding support is given in grant-funded research projects

It should also be recognised that many agencies within government have struggled to find the resources required to make their own data available in a timely and rigorous fashion. This means that data becomes available too late to affect the policy process, or is made available in a way that leads to error and less than robust conclusions. To implement the recommendations within the draft report, it will be necessary to

ensure not only an adequate level of resources for the university sector, but also those data providers who the university sector interacts with.

These costs should be recognised as an investment. Making data available in a robust, timely and safe way will make it less likely that government expenditure is poorly directed or inefficient, and also identify areas where government expenditure can have positive economic returns.

**Goal 5. Recognition that there are some communities and groups within society for whom data ownership and control is particularly important**

While the ANU is strongly in support of the default stipulated in the Productivity Commission draft report that de-identified data be made available for re-use and analysis, it also needs to be recognised that certain types of data collections may require additional levels of control. In particular, there is a strong need to ensure that data that is collected on or by Aboriginal and Torres Strait Islander communities is made available in a way that protects the rights and ownership of that data. That does not necessarily mean that such data is not made available to other researchers from outside the community. Rather, it means that the communities have sufficient control over access, and are sufficiently resourced to ensure that the data is used in a way that directly benefits the community.

Please don't hesitate to contact my Executive Officer, Karen Warnes ([eo.dvcr@anu.edu.au](mailto:eo.dvcr@anu.edu.au)), should you require further information or clarification of the points above.

Yours sincerely

A handwritten signature in black ink that reads "MM Harding". The signature is written in a cursive, flowing style.

Professor Margaret Harding  
Deputy Vice-Chancellor (Research)