29 July 2016

Data Availability and Use
Productivity Commission
GPO Box 1428

RE: Inquiry into Data Availability and Use

Statistics NZ welcomes the opportunity to make a submission on the Issues Paper on Data Availability and Use.

Statistics NZ is the New Zealand government department responsible for the collection and production of official statistics and data. Statistics NZ's purpose is to empower decisions by adding value to New Zealand's most important data. Our goal is to increase the value of data to decision-makers 10-fold in the next 15 years. This will require focused and coordinated effort across the NZ data ecosystem – and how data is made available and used plays a central role in this.

Data Availability and Use in New Zealand – the Integrated Data Infrastructure

As you outlined in the inquiry's Issues Paper, the amount of data being produced, collected and therefore available for potential analysis has grown rapidly in recent years. Our submission will focus on New Zealand's Integrated Data Infrastructure (IDI), a world leading approach to making anonymised, integrated data safely available for research and analysis, to inform evidence-based decision-making and policy.

There is an increased focus on evidence-based decision-making in New Zealand government practice, principally through an investment approach to government spending (with an initial focus on the social sector) which aims to improve the lives of New Zealanders. The approach requires the use of evidence prior to investment decisions being made and it also aims to increase the measurability of the impact of government investment. The IDI is the key data repository which provides the evidence base for this approach.

The IDI has begun to attract international interest, most recently from the European Union (EU). The NZ Government Statistician was invited to speak to EU commissioners and others in June 2016. This led to invitations to speak to various other European conferences as well as an explicit interest from EU researchers to come to New Zealand to learn more about the IDI, and how it works.

A short history of data integration and the IDI

Statistics NZ has been undertaking data integration since 1997. The early integration projects were for specific purposes, often linking two or three datasets, with each kept in a separate environment – the first ones linked employer and employee data, employment outcomes to tertiary education data, and Census data to the NZ cancer mortality register.

Our early partners in data linking often had expertise in administrative datasets and were interested in expanding their current analytical sphere by adding their administrative data to survey data or other administrative data.

In 2011, Cabinet agreed to a proposal to integrate Department of Labour migration data with data in linking projects managed by Statistics NZ – this resulted in the establishment of the IDI.

Support from Government grew as the system's data needs were better understood and integrated data began to deliver first results. In 2013, Cabinet agreed that delivery of the Better Public Services results[1] would benefit from improved capability across government to share data using existing datasets. A cross-agency data-sharing solution was required. As a way forward, Cabinet agreed to expand the IDI to facilitate this work, and to Statistics NZ (building on its earlier data integration work) housing the IDI. For the Cabinet paper on this decision, see here: http://bit.ly/2afUqJa.

How the IDI works – data integration and privacy go alongside each other

A key aspect of the IDI is that it has multiple security layers in place to ensure the data is not misused. Data is anonymised at the start and confidentialised at the end. Access to the IDI only occurs under stringent conditions and Statistics NZ's 'five safes' framework governs the access to our microdata, including to the IDI: safe people, safe projects, safe settings, safe data, and safe output – for more detail, see here: http://bit.ly/2awIHFL.

Social licence – public generally supportive of data use for public good

Because the IDI is linking and using data to an extent previously unseen, the question of social licence is a crucial one. Recent research on social licence and perceptions about data collection and use showed the New Zealand public has an expectation that their data is used, but used wisely and for public good. It also showed that Statistics NZ has strong reputation as 'data expert' and trusted custodian of public information. However, this should not be taken for granted – social licence, once earned, can be easily lost.

Statistics NZ is therefore continuing to engage the public in an informed debate about social licence, intended to move them from naïve to informed trust. We are collaborating on this with the NZ Data Futures Partnership, a cross-sector group set up to help drive change across New Zealand's data-use ecosystem, as they are leading the public conversation with New Zealanders about social licence.

Insights and impacts from the IDI

Researchers from across government and academia have been using the IDI to answer a wide range of research, policy, and evaluation questions. Previously unanswerable questions can be studied using the IDI because of the richness of linked data, and support the approach of the NZ Government to drive greater availability and use of data for the public good.

Ultimately, the IDI helps the Government and its agencies to prioritise and maximise the impact of expenditure. IDI data can be used in liability modelling, to reduce current and future liability in areas of big government spending, such as the social sector.

A list of IDI research projects can be accessed here: http://bit.ly/2awIHFL, and some notable examples which demonstrate the value to be derived from increased data access are outlined below.

IDI research supports the investment approach for at-risk families

The new investment approach for at-risk families involves gaining a better understanding of the situations of vulnerable children and their families, and the factors that lead to poor outcomes later in

[1] In 2013, the Government set ten results for the public sector to achieve over the next five years. Delivering these Better Public Services results within tight financial constraints is one of the Government's four priorities (the others being responsibly managing the Government's finances, building a more competitive and productive economy, and supporting the rebuild of Christchurch).

life. The Government is currently modernising Child, Youth and Family (CYF) services by applying the social investment approach to at-risk families and vulnerable children. It has set up a Social Investment Unit (SIU) to lead this change.

The Treasury's Analytics and Insights team used the IDI to identify four key indicators for children at high risk of poor outcomes later in life. This work aims to identify where to invest earlier into vulnerable children and their families, rather than deal with problems after they have emerged. The work is shared with the community via an interactive tool: https://shinyapps.stats.govt.nz/sii/.

The IDI supported the development of the NZ Government's Youth Funding Review and Vulnerable Children's Plan which outlines planned investment for better outcomes for youth and links and monitors services. Key to this approach is working across sectors and the development of operational insights. Operational information sharing, however, occurs outside the IDI.

IDI research supports better crime information

The investment approach to justice is a cross-sector programme of work that is managed through the Justice sector group and is supported by multiple agencies. Early research done through the IDI found that after the age of 22, most opportunities to prevent offending are lost. To invest in the right place and at the right time, risk assessments are crucial to focus on the few who will have long offending careers. The research also showed that many crime prevention opportunities occur before the first offence and before 14 years of age.

Following this research, the Justice sector is now using the IDI to develop a full actuarial approach. For more information, see here: http://bit.ly/2aLzEgl.

IDI research supports better health outcomes

The IDI represents great potential to monitor health treatments on a large scale and consider their effectiveness. The New Zealand integrated data model is of high interest to epidemiologists worldwide as the IDI contains information on a range of conditions, diagnoses, and treatments. It is seen to have particular value in understanding the social determinants of health. A Virtual Health Information Network has been formed to co-ordinate and optimise work among New Zealand researchers.

IDI research supports economic analysis

The Longitudinal Business Database (LBD) complements the IDI with microdata about businesses. It includes tax data and business survey data and has been used to produce findings on:

- the impact of research and development subsidies on innovation: a study of New Zealand firms' foreign acquisitions and the performance of New Zealand firms
- labour market dynamics following a regional disaster (the Christchurch earthquakes)
- insights into factors impacting New Zealand's productivity.

The IDI today

Today, there is widespread support for the IDI and Statistics NZ's management of it. The IDI has developed much quicker and much further than originally envisaged, with more datasets than originally anticipated already added, and continuing to be added – not only from government agencies but also NGOs, iwi (indigenous tribes), and the private sector.

In fact, we have moved from asking agencies for permission to add their data to agencies requesting that their data be added. The value of the insights produced through the IDI is now generating increased demand in its own right and there is an increasing expectation from Ministers that agencies use the IDI for their investment decisions and bids.

We are undertaking significant efforts to boost access to IDI data to better meet demand, including an increased number of remote datalabs (16 datalabs as at June 2016) and improvements to the quality of the technical infrastructure.

Capacity to use the data has developed quickly but not as fast as demand has. The size of the data analyst community has grown and current IDI use goes beyond those envisaged: it is becoming an important epidemiology tool, it has more utility for service delivery evaluation than was initially predicted, and there is more interest from non-government data providers than we expected.

Our current legislative settings enable and protect us in doing data integration, and a current review of the Statistics Act 1975 will ensure that we can continue, and, potentially, within our social licence mandate, expand this work in the future.

Lessons learnt

The IDI story so far has been one of significant success and has provided Statistics NZ with some important learnings which could be helpful pointers when considering integrated data use and access elsewhere:

- The backing by Ministers provided us with valuable support in expanding the IDI – the fact that the IDI enables the recently adopted investment approach has made it an integral part of the institutional setting for policy and evaluation in New Zealand.

- The modular infrastructure of the IDI has allowed us to add more data over time and this has worked well – however, while it was the right way to set up the IDI at the time, technology is rapidly improving and there are better solutions available now – this points to a need to continue investing in data access mechanisms, including those for non-technical users, beyond the start-up phase.

- The existence of the IDI has made us and other agencies more aware of the immense value of the government's data asset, and we have added more data to the IDI than we were originally funded for (at no additional cost) – once there was a place for data integration and the benefits became visible, need and demand quickly followed.

The IDI has revolutionised the availability and use of data in New Zealand and has become an integral part of policy and decision-making – its full potential is yet to be realised. Statistics NZ also eagerly awaits similar data linking infrastructures being set up elsewhere in the world, so we can continue to share our expertise in this area but also learn from others and their experiences.

Yours sincerely,

Liz MacPherson
Government Statistician and Chief Executive