Received: 22 August 2023 | Revised: 9 October 2023 | Accepted: 16 October 2023

DOI: 10.1002/ajs4.300

COMMENTARY

WILEY

Data and evaluation: A match made in policy heaven

Andrew Leigh

Parliament of Australia, Canberra, ACT, Australia

Correspondence
Andrew Leigh, Parliament of Australia, Canberra, ACT 2600, Australia.
Email: andrew.leigh.mp@aph.gov.au

Abstract

Better data are fundamental to improving the effectiveness of policies. Drawing on examples from nutrition, education, criminal justice and income support, I discuss how data access has contributed to policy improvements. Data linkage across departments can also help ensure that policymakers are focussed on the right policy goals, rather than merely those that are easiest to measure. As governments increase the rigour of evaluation, quality data will be crucial.

KEYWORDS

economic policy, employment policy, health economics, public administration, randomised trials, social policy

1 | INTRODUCTION

Quality data are fundamental to understanding social disadvantage. Additionally, as governments craft policies to improve equity, good data are essential for policy evaluations. Data therefore play a dual role in the policy development process—shining a light on potential problems and helping judge the effectiveness of potential solutions.

In this article, I discuss how data quality and data access have contributed to policy improvements, drawing examples from nutrition, education, criminal justice and income support. Major data linkage projects have improved the microdata available to government agencies and academic researchers. These new administrative datasets not only make it possible to explore the relationship between different markers of disadvantage, but also enable evaluations to be conducted more cheaply.

Evaluation is critical to the operation of government. When subject to rigorous evaluations, US studies find that many as nine in 10 education programmes, employment programmes and new pharmaceuticals fail to produce a measurable positive impact

This is an edited version of a speech delivered at the Data for Policy Summit, organised by the Life Course Centre, in Canberra on 15 August 2023. My thanks to officials in the Australian Treasury for drafting assistance, and to an anonymous referee for valuable feedback.

(Leigh, 2023). As sociologist Peter Rossi has noted, "the better designed the impact assessment of a social program, the more likely is the resulting estimate of net impact to be zero" (Rossi, 1987).

The establishment in 2023 of the Australian Centre for Evaluation (ACE) marks a shift in the way that the federal government views evaluation. Quality data will be critical to ACE's work, allowing rigorous evaluations, including randomised trials, to be conducted in a more cost-effective manner than would have been possible if evaluators relied on survey data alone.

The remainder of this paper is structured as follows. Section 2 provides a historical example, from the field of nutrition, in which data helped to shape policy. Section 3 discusses a trio of contemporary examples: schooling policy, incarceration policy and anti-poverty policy. Section 4 discusses new innovations: data linkage to support employment services reform, the Australian Centre for Evaluation and new integrated datasets. In this section, I also discuss concerns of security, privacy and ethics. The final section concludes.

2 | DATA HUNGRY

The story of using data to inform policy has a long lineage.

In the mid-1930s, the Australian Government appointed a Commonwealth Advisory Council of Nutrition to undertake the first national survey of nutrition.

In the midst of the Great Depression, the Council put together a team of health and agricultural experts including scientists, researchers, doctors and dentists. State committees also played a part.

The Council conducted two significant studies.

The first was initiated via a daily questionnaire published in capital city newspapers asking people about the food they purchased and the food they ate that day. Some participants were invited to continue in the study for up to a year.

The Council ended up analysing the food records of around 1800 households as part of the "dietaries" study (Senate Standing Committee on Social Welfare, 1979).

It found Australians of the 1930s were "generally well fed," but questions have since been raised about whether the sample was truly representative of the population (Senate Standing Committee on Social Welfare, 1979).

For the second study, the Council sent a doctor out to examine almost 6000 children as part of "a survey of the nutritional state of children in inland areas" (Australian Bureau of Statistics (ABS), 1940).

The results of this survey were more concerning. It found "a considerable mass" of malnutrition—between 13 per cent in rural Victoria and up to 24 per cent in New South Wales (O'Connell, n.d.).

The findings from these studies led the Council to recommend the formation of a child health division in the Department of Health.

It also called for medical supervision in kindergarten and primary schools nationally as well as reduced milk prices.

School milk programmes have their genesis in this data collection process.

In the modern era, the linkage between data and policy continues. The National Health and Medical Research Council is currently working with experts to update the Australian Dietary Guidelines. This is based on surveys showing that only one in 20 adults and one in 17 children met both the fruit and the vegetable daily recommendations (Australian Bureau of Statistics, 2018).

3 | DATA DELIVERS

If data and evaluation are a match made in heaven, everyone would agree that the stairway to get there can take some climbing. Consider three examples—from schooling, imprisonment and income support.

3.1 | Schools learn from each other

In 1999, the Ministerial Council on Education, Employment, Training and Youth Affairs released the Adelaide Declaration on National Goals for Schooling in the 21st Century—an agreement to progress toward a higher national standard for schooling, including reporting nationally on school achievement.

It was a long road but more than a decade later in January 2010, the MySchool results went live for the first time. This allowed parents to see how their child's school or prospective school stacked up against other similar schools. It also allowed school administrators to assess their relative strengths and weaknesses.

Over the years, the MySchool reporting has encouraged schools doing well to continue their good work. For schools that are looking to improve, they can more readily identify exemplars that they can learn from.

Some of those schools shared their inspirations and lessons learnt (Australian Curriculum, Assessment and Reporting Authority (ACARA), 2017). One principal said their school's success story was all about data—the school was "unaware of severe downward trends in literacy and numeracy that had been occurring over many years." Another principal said the school identified gaps and restructured intervention programmes. Others invested in staff programmes and professional learning.

3.2 | Closing prisons

Another successful example of data-driven policy comes from the United States and involves addressing concerns about mass incarceration and the need for criminal justice reform (Bureau of Justice Assistance (BJA), 2012).

For 3 years, from 2006 to 2008, the United States incarcerated more than 1 per cent of adults (Leigh, 2020). The human and economic cost of this was enormous.

Launched in 2010, the Justice Reinvestment Initiative used criminal justice data to identify areas of overincarceration, assess offender risk and estimate future prison populations.

By understanding these trends, US policymakers could tailor interventions. They could put in place diversion programmes for low-risk offenders, and they could invest in community-based rehabilitation and treatment services.

The adoption of the Initiative was associated with significant reductions in prison populations in various states. These states reduced correctional costs while maintaining public safety.

For example, Utah used funds to identify alternative justice pathways that proved successful in the past. Increased use of these pathways reduced the use of incarceration and directed people to options that decreased the level of recidivism (Bureau of Justice Assistance (BJA), 2012).

Oregon took a different direction. It undertook a system analysis of the needs of people in the criminal justice system. The state eliminated expensive duplication of services while also reducing the reliance on prison as a means of behavioural health treatment (Bureau of Justice Assistance (BJA), 2012).

During the 2010s, the United States experienced a drop in crime and a fall in the incarceration rate. In Republican-run and Democratic-run states, prisons were closed. Even with the increase in crime that occurred during the COVID-19 pandemic, rates of violent crime and rates of imprisonment are substantially lower than they were 15 years ago. Data and evaluation have played a critical role.

3.3 | Kiwi data

Integrated Data Infrastructure—the New Zealand Government's data tool—is also a data success story with more than 730 projects using its microdata in the past decade (Song, 2022).

In research on social investment in New Zealand, Janeen Baxter and Sara Kalucza noted:

> While the Integrated Data Infrastructure was established and is run independently of the Social Investment Agency, it was fundamentally a part of the social investment approach to integration of data into policy, and achieving the key goal of identifying and tracking population cohorts over time and the opportunity to evaluate interventions or system-wide policy changes.
>
> (Kalucza & Baxter, 2021)

Integrated Data Infrastructure works by linking or marrying up administrative data from multiple government agencies. Data cover health, justice, labor market, social development, housing and education—and in some cases going back to 1840.

This enables researchers and policymakers to study individuals and populations over time. It also facilitates longitudinal analysis that provides insights into social trends, life events and policy impacts.

All identifying variables are removed to protect the individual's privacy prior to release. These data also link individuals via taxation data to the Longitudinal Business Database, a microdata resource about New Zealand businesses.

The comprehensive nature of the Integrated Data Infrastructure system presents many opportunities for policymakers to investigate policy impacts across most portfolios in the New Zealand Government. For example, where people are located, how many there are and how great their need is.

In one example, a researcher from the NZ Treasury used a combination of data to examine how health conditions such as diabetes, stroke and cancer can affect people's income and employment (Dixon, 2015).

In another example, the Ministry for Children (Oranga Tamariki) used the Integrated Data Infrastructure to better understand which children in which locations were at risk of ending up in the foster care system (McCann, 2019). Efforts were targeted at the locations identified through the dataset.

4 | BIG DATA OPPORTUNITIES

Like all relationships, data and evaluation can always improve and work better in tandem. In this section, I discuss three examples—employment services reform, the Australian Centre for Evaluation and integrated datasets—and review some of the major challenges facing integrated data and rigorous evaluation.

4.1 | Employment services

The government's employment services programme, known as Workforce Australia, provides an example.

The programme's main objective is to help job seekers find and maintain secure work. Therefore, it is reasonable to measure the success of the programme on whether it enhances the chance of a job seeker finding a job and remaining employed.

We know from job seeker surveys and employment administrative data whether programme participants are employed, but these data do not capture everyone, nor do they always capture the nature of their employment.

The Department of Employment and Workplace Relations evaluates the impact of the programme using exit measures. For example, the Department uses exit from employment services and from income support as a proxy for employment, and it uses the duration of the exit as a proxy for employment sustainability.

The challenge for policymakers is these exit measures tell us nothing about job *quality*.

Studies tell us that the level of earnings is not the only factor that determines the job quality, but it is undoubtedly an important factor.

A study on the quality of employment among young Britons found no universally satisfactory definition of job quality (Orlando, 2021). However, it found common traits that include a good and fair income, job security and stability, opportunity to progress, a good work-life balance and one that gives employees a voice.

A US study—titled *Not Just A Job: New evidence on the quality of work*—found that 60 per cent of survey respondents self-assessed as having a mediocre or bad job (Rothwell & Crabtree, 2019). This study's list of what constitutes a good job looks very similar to the UK one: pay, job security, opportunity for advancement, benefits, stability and dignity.

The OECD also uses earning quality as one of the three ways to measure and assess job quality.

A decent life needs a decent income, and a decent level of income comes from decent earnings. Therefore, where possible, earnings should be included as an outcome measure in evaluating the employment services programme.

The inclusion of earnings in evaluating employment services programmes can shift the mindset about how these programmes are expected to work.

There's an opportunity to build on the "work-first" approach that most of the current programmes are taking with an aim to moving job seekers off income support.

Using earnings as an outcome allows us to devote more attention and resources to the development of job seekers' human capital, which is the most important factor determining earnings.

For these reasons, earnings data should be added to the data assets held by the department. This is about enhancing their capability to accurately measure job seekers' labour outcomes, and consequently to conduct more meaningful evaluations of the programme's impacts.

The Department of Employment and Workplace Relations is working with the Australian Taxation Office and Treasury to explore options to integrate the tax office's earnings data into employment services data.[1]

Linking tax data to the unemployment system would only provide information on pay—it would not say much about security, stability, progression, balance, voice or dignity.

However, including tax data could be a significant first step toward understanding the level of employment a person has achieved.

In the longer run, better data will improve the quality of the Department's evaluation work.

4.2 | The Australian Centre for Evaluation

More broadly, the establishment of the Australian Centre for Evaluation provides an opportunity for evaluation to improve public programmes and save taxpayer money.

The Centre will promote the use of high-quality evaluation across the Australian Public Service. It will partner with government agencies to initiate a small number of evaluations each year. It will work to improve evaluation capabilities, practices and culture across government.

In particular, the Centre will implement multiple randomised control trials, focussing not just on large trials, but also on quick and economical experiments.

One example of this occurred in 2014 when the Obama Administration pushed for better evidence-based policy via low-cost randomised control trials.

In some cases, trials can be low cost because the administrative data have already been collected. In other cases, the intervention can be extremely low cost.

For example, a child substance abuse programme in Illinois cost about $100,000 over 10 years, but saved around 60 times that much through shorter and fewer foster placements (Coalition for Evidence-Based Policy, 2012).

In another low-cost example, a New York incentive programme paid teachers in low-performing schools a bonus if results improved. The school board already collected the outcome measures, so the bonus offered to teachers was the only expense. In this case, the administrative data clearly showed the lack of effect of the bonus payments (Coalition for Evidence-Based Policy, 2014).

We have seen low-cost randomised control trials in Australia. In 2016, Rebecca Goldstein and Michael J. Hiscox led a trial to evaluate changes in the School Attendance and Enrolment Measure programme in the Northern Territory (Goldstein & Hiscox, 2018).

The programme was a type of "welfare sanction." In other words, participants could lose income support payments if they failed to comply with conditions—in this case, failing to address a poor school attendance record.

The outlay for this randomised control trial was minimal. It required no materials, no travel and a short time for the researchers to pair like students for the treatment and control groups.

The researchers looked at attendance records for both groups and found "no significant differences" following any of the programme interventions. As a result of this evidence, the programme was subsequently closed.

4.3 | Integrated datasets

Being able to use data for policy development and efficient service delivery is critical to the operation of the public service.

The Australian Bureau of Statistics is doing its part to enhance the value of their data. This will improve the capability of the government to assess need, efficiency and effectiveness.

As chief statistician David Gruen has said "while there is always more to do, we are making excellent progress improving the quality and timeliness of administrative data and, more broadly, integrated data assets" (Gruen, 2023).

Another priority at the Australian Bureau of Statistics, right behind continuing to produce high-quality primary stats about Australia, is the desire to make secondary products—such as the Business Longitudinal Analysis Data Environment (BLADE) and the Person-Level Integrated Data Asset (PLIDA)—more useful and informative for understanding ourselves.

The government is committed to ensuring the public service has the right capability and tools. Ultimately, this will lead to better policy advice, better regulation and better services.

4.4 | Challenges

The development of integrated data assets—and the use of big data in policy more broadly—depends on ensuring that data are kept secure and privacy is maintained. With cyberattacks affecting large firms such as Optus, Medibank and Latitude Financial Services, custodians of large data assets need to continuously monitor for intrusions into their systems, and test for vulnerabilities in hardware, software and "liveware" (personnel).

Individual privacy must also be protected. In 2017, Coalition Human Services Minister Alan Tudge and his department provided the media with personal details of the welfare history of a commentator who had criticised government policy. Such a misuse of personal data has the potential to undermine public confidence in integrated data. In 2023, the Australian Government announced an expansion of the Office of the Australian Information Commissioner, including the appointment of a dedicated privacy commissioner.

Rigorous evaluation faces challenges of its own. If government programmes turn out to be ineffective, then governments need to be willing to acknowledge the findings, and adapt accordingly. An "experimenting society" requires a measure of modesty, and an intellectual rigour that seeks continuous improvement. Randomised evaluations face challenges of their own, with 27 per cent of Australian politicians (Ames & Wilson, 2016) and 17 per cent of Australian voters (Biddle et al., 2023) opposing the use of randomised trials. Randomistas have more work to do in persuading the sceptics.

A common concern about rigorous evaluations is over ethics, with some critics arguing that it is unethical for people to randomly miss out on a potentially beneficial government programme. However, failing to evaluate programmes has ethical issues of its own. Is it really ethical to spend scarce taxpayer dollars on programmes that may be ineffective? An expansion of randomised trials will only be successful if the ethical review framework is as rigorous as the evaluation methodology itself.

5 | CONCLUSION

More than four decades ago, a Parliamentary Committee inquiry into evaluation in Australian health and welfare services noted:

> The principal value of evaluation is that it improves the decision-making process by providing a rational base from which judgements may be made.
>
> (Senate Standing Committee on Social Welfare, 1979)

The Committee noted, "Without evaluation we cannot know whether a particular program is achieving anything at all or whether, for example, its effects are the reverse of its stated objectives." The first chapter of the Committee report is devoted to what evaluation is, and why evaluation is desirable. It describes the four essentials for any evaluation as: a statement of needs; a statement of objectives or goals; a statement of the criteria by which success will be judged; and a database of evidence to measure these criteria.

The Committee went on to make a series of recommendations to improve data, develop indicators and spend more on health and welfare statistics. The final recommendation sounds a bit like the Australian Centre for Evaluation's objectives. The Committee called for a Secretariat to prepare a document outlining the methods available to organisations for the evaluation of their activities.

The Committee also recommended that the Department of Health and Social Security provide a consultancy service, free of charge, to enable organisations receiving health and welfare grants from the Federal Government to evaluate their own activities.

Over the past half century, some things have remained constant. Data and improved evaluations have been a long-standing issue and a vital one. However, what has changed is the data and digital environment—the capacity to generate, synthesise, analyse and share data analytics at scale, and in real time, has fundamentally shifted.

Funnily enough, the 1979 Senate Committee's report was titled *Through a Glass, Darkly* in referring to Saint Paul's point that we cannot see clearly now but if we get things right all will be clearer in heaven (I Corinthians 13:12).

Data and evaluation really are a match made in policy heaven.

AUTHOR CONTRIBUTIONS

Andrew Leigh: Conceptualization; writing – review and editing.

ORCID

Andrew Leigh https://orcid.org/0000-0002-5639-0509

ENDNOTE

[1] Note that while employment and taxation data are already available in the Person-Level Integrated Data Asset (PLIDA), this dataset is designed to preserve individuals' anonymity. By contrast, data linkage between the employment services system and the taxation system is designed to provide the Department of Employment and Workplace Relations with a more comprehensive picture of the well-being of individual jobseekers.

REFERENCES

Ames, P. & Wilson, J. (2016) *Unleashing the potential of randomised controlled trials in Australian governments*. M-RCBG Associate Working Paper No. 55, Cambridge, MA: Mossavar Rahmani Center for Business & Government, Harvard Kennedy School.

Australian Bureau of Statistics (ABS). (1940) Year Book Australia, 1939 Chapter IX, Public Hygiene, p. 222.

Australian Bureau of Statistics (ABS). (2018) National Health Survey: first results [reference period 2017–18 financial year], released 12 December 2018.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2017) *School perspectives: principals of schools achieving substantially above average gain in NAPLAN 2017*. Sydney: ACARA.

Biddle, N., Gray, M. & Hiscox, M. (2023) *Public support for randomised controlled trials and nudge interventions in Australian Social Policy*. Working Paper, Canberra: Centre for Social Research and Methods, Australian National University.

Bureau of Justice Assistance (BJA). (2012) Justice reinvestment initiative.

Coalition for Evidence-Based Policy. (2012) *Rigorous program evaluations on a budget: how low-cost randomized controlled trials are possible in many areas of social policy*. Washington, DC: Coalition for Evidence-Based Policy.

Coalition for Evidence-Based Policy. (2014) *Policy New York City teacher incentive program*. Washington, DC: Coalition for Evidence-Based Policy.

Dixon, S. (2015) *The employment and income effects of eight chronic and acute health conditions Treasury*. Wellington: New Zealand Government.

Goldstein, R. & Hiscox, M. (2018) *School enrolment and attendance measure randomized controlled trial*. Canberra: Full Report National Indigenous Australians Agency.

Gruen, D. (2023) Making the most of administrative and integrated data assets [speech delivered 3 August 2023 Australian Workshop on Public Finance].

Kalucza, S. & Baxter, J. (2021) Social investment: the New Zealand case. In: Chan, J. & Saunders, P. (Eds.) *Big data for Australian social policy: developments, benefits and risks*. Canberra: Academy of the Social Sciences in Australia, pp. 24–37.

Leigh, A. (2020) Estimating long-run incarceration rates for Australia, Canada, England and Wales, New Zealand, and the United States. *Australian Economic History Review*, 60(2), 148–185.

Leigh, A. (2023) Evaluating policy impact: working out what works. *Australian Economic Review*, forthcoming.

McCann, D. (2019) *Experiences of education for children in care: part2: review of New Zealand Government Data*. Wellington, New Zealand: Oranga Tamariki—Ministry for Children.

O'Connell, J. (n.d.) Commonwealth Advisory Council on Nutrition. Australian Food Timeline. [Accessed August 2023].

Orlando, C. (2021) *'Not just any job, good jobs!': youth voices from across the UK*. Brighton: Institute for Employment Studies.
Rossi, P.H. (1987) The iron law of evaluation and other metallic rules. *Research in Social Problems and Public Policy*, 4, 3–20.
Rothwell, J. & Crabtree, S. (2019) *Not just a job: new evidence on the quality of work in the United States*. Indianapolis: Lumina Foundation.
Senate Standing Committee on Social Welfare. (1979) *Through a glass, darkly: evaluation in Australian Health and Welfare Services*' Volume One, Canberra: Australian Government Publishing Service.
Song, Z. (2022) *How big data can be a force for good – NZIER Insight 103*. Wellington: New Zealand Institute of Economic Research.

AUTHOR BIOGRAPHY

Andrew Leigh is a Assistant Minister for Competition, Charities, Treasury and Employment, and the Federal Member for Fenner in the Australian Parliament. Prior to being elected, he was a professor of economics at the Australian National University. His books include Randomistas (2018), Reconnected (2020) and Fair Game (2022). Andrew is a keen triathlete and marathoner, and lives in Canberra with his wife Gweneth and their three sons. His website is andrewleigh.com.

How to cite this article: Leigh, A. (2023) Data and evaluation: A match made in policy heaven. *Australian Journal of Social Issues*, 00, 1–9. Available from: https://doi.org/10.1002/ajs4.300