

Wicked problems: using data for better public policy

PBO Conference

Danielle Wood, CEO, Grattan Institute 3 November 2022



A golden age for data driven policy analysis?

- Linked administrative data is a game changer
- Large proprietary data sets will also be significant
- Both allow us to answer new questions and to form more timely policy responses

There are still barriers in using data to its full potential

- Data is still expensive, slow to arrive and sometimes difficult to access
- More expertise is required within the public service
- Poor quality 'black box' data analysis in contentious policy debates

Unleashing the golden age

- Boosting data culture embrace need for data availability and timeliness not just for admin data
- Boosting capability within all agencies to use the data
- Framework for access to proprietary data sets
- Countering poor quality analysis



Data, data everywhere

Data volume created and replicated worldwide (in zettabytes)







Chief Statistician ABS, David Gruen

It seems clear that the potential value of data increases many-fold when individual data sources are brought together to enable public policy issues to be examined from a range of different perspectives. Multi-Agency Data Integration Project (MADIP) provides data on individuals and can GRATTAN be linked across a huge number of policy areas





Similarly for business....

Business Longitudinal Analysis Data Environment (BLADE)

What is **BLADE**?

- Spine contains all businesses (ABNs) that have been active at any time between 2001/02 and 2018/19
 - Sourced from the ABS Business Register
 - Updated Annually
- BLADE brings together:
 - Taxation Information
 - ABS Survey Data
 - Economic Activity data such as Imports and Exports
 - Intellectual Property information
 - Government Program Administrative data



Also

- A-Life (ATO)
- Single Touch Payroll (ATO)
- NIHSI (AIHW)

The questions are endless.....





Research using MADIP/BLADE – a flavour

- Mental health following the 2019-20 bushfires: service use, help-seeking behaviours and barriers to access in bushfire affected communities (2021) -PM&C
- Women in STEM analytical project (2020) DISER
- Reducing life lost in heatwaves (2019) Bureau of Meteorology
- Competition in Australia and its impact on productivity growth (2020) The Treasury
- Intergenerational drivers of entry into the Victorian Child Protection system (2019) – Victorian Department of Health and Human Services
- Development of an evidence-based NSW Early Childhood Education Outcomes Framework (2020) – NSW Department of Education
- Employment effects of the JobKeeper program (2020) ANU
- Predicting the Value and Use of Urban Land (2019) University of Queensland
- Age Pension Income Means-Test and Late Life Labour Supply (2019) University of Sydney





Example: Reforming temporary skilled migration







Current approach to rationing temporary skilled migration is to use occupation skill lists



Each dot represents 1,000 full-time jobs in an occupation by their income



III-time workers aged 19 and older in the 2016 Census. Skills list is the STSOL. Competency scores are the average of 10 core occupational competencies – such as numeracy and problem solving – developed by the National Skills Commission.

Occupations are a poor way to target high wage jobs: nominated incomes vary dramatically *within* occupations





Source: Grattan analysis of ABS MADIP (2021).

11

Proposal: award points based on income rather than an arbitrary 'skills list'



Each dot represents 1,000 full-time jobs in an occupation by their income



A \$70,000 wage threshold would exclude about 30 per cent of TSS visas, a \$90,000 C threshold would exclude about 60 per cent



Cumulative share of TSS visa grants by nominated income



Note: TSS visas awarded in 2018. Source: Grattan Analysis of ABS (2021).

Hospitality would be most affected losing about 1.5% of its workforce, other sectors less affected





This would also yield a fiscal dividend



Annual fiscal benefit of temporary skilled primary applicants, in 2018 and with different settings



Grattan analysis of MADIP and Varela et al (2021).

Proprietary data sets are increasing valuable in dollar terms but also for understanding policy questions





https://howmuch.net/articles/biggest-brands-of-the-world-2020

For example, During COVID proprietary data sets provided real time data on economic behavior and movement





Source: AlphaBeta, Illion credit and debit card data



A golden age for data driven policy analysis?

- Linked administrative data is a game changer
- Large proprietary data sets will also be significant
- Both allow us to answer new questions and to form more timely policy responses

There are still barriers in using data to its full potential

- Data is still expensive, slow to arrive and sometimes difficult to access
- More expertise is required within the public service
- Poor quality 'black box' data analysis in contentious policy debates

Unleashing the golden age

- Boosting data culture embrace need for data availability and timeliness not just for admin data
- Boosting capability within all agencies to use the data
- Framework for access to proprietary data sets
- Countering poor quality analysis

Challenge: improving access to microdata



- A project can take many months to get approved – a problem for timely policy analysis, also very difficult for project management and planning
- Available variables for custom extracts are not published, making project scoping challenging
- Mandatory training opportunities are infrequent
- Data is very costly (e.g. about \$9,000 for a recent project on health care out-ofpockets). This a barrier especially for NGOs but also academics
- Difficulty in qualifying as a trusted user if outside academia

Grattan examples:

Rental holiday for businesses in hibernation during COVID

- Applied March 2020 (under urgent COVID research stream)
- Quick approval: 2-3 weeks
- 10 months to offer compulsory training need for data access

Temporary skilled migration project

- Applied October 2021
- Approvals February 2022 (held up by one agency)



Some improvements have been made but still....

- The secure platform regularly crashes
- Data files are so big, they can take hours or days to process
- No ability to resolve dataset quirks as assistance from data owners (e.g., relevant departmental people) is restricted due to privacy constraints







PUBLIC SECTOR DATA MANAGEMENT

....there is a global undersupply of data and analytics which limits the ability to get the most value out of publicly available data. Ready for work graduates with data capabilities are in short supply. During consultations most agencies expressed a need for more data capabilities

July 2015



This is not just a government problem.....

Share of businesses who use data analytics (2019) and artificial intelligence (2020)

Data analytics

Artificial intelligence



Source: Productivity Commission, 5 Year Productivity Inquiry: Australia's data and digital dividend (2022, Figure 1.12).

Advanced analytical methods are increasing in Australia, but slowly





Figure 3: Share of job ads requiring technological skills

Notes: 12-month rolling average of monthly share of job ads that require each technology skill. These charts use Lightcast's predefined skill categories, rather than the text analysis approached used in Figure 2. Cloud Computing also includes Cloud Storage and Cloud Solutions. Source: Treasury analysis of Lightcast data for Australia and the United States (2012-2020).

Source: Bahar and Lane (2022, p. 24) Treasury: https://treasury.gov.au/sites/default/files/2022-10/p2022-325290-full-report.pdf

Challenge: proliferation of self-interested 'black box' data analysis in contentious policy debates





Q7 tl 1 0 2 1

Rents to rise, house prices to decline under Labor's negative gearing policy: SQM

Domain March 20, 2019

Property faces a 'day of reckoning' if Labor scraps negative gearing hurting every Australian

News.com.au 4 April 2019

Negative gearing policy a \$1.5 billion voter turn-off

The Australian 19 August 2019

A lot of it is terrible....





- BIS Shrapnel uses CGE model to estimate winding back of negative gearing would reduce GDP by \$19 billion a year
- Simply not plausible that a **\$2 billion tax change** drives economic hit so large

Media will run numbers (particularly big ones) with no critical lens

\$19bn hit in Labor negative gearing strike on investors: report

The Australian, 3 March 2016 (page 1)

Negative gearing changes would push up rents 10 per cent: BIS Shrapnel

Domain, 3 March 2016

Negative gearing: Modelling reveals it could push up average rents by \$2600 a year

Daily Telegraph, 3 March 2016

o Bill's	s gearir	ig bio
Pure		Jan Start St

gative outlook



 Sector Biology and an estimated mean and an and mean sector bases before the sector sector and sector bases before the sector sector and sector bases before the sector bases before the sector sector bases before the sector sector bases before the sector bas sector bases before the sector bases

Herald Sun, 2 March 2016

PROGNOSIS NEGATIVE

BIS SHRAPNEL SAYS THAT IF NEGATIVE GEARING IS LIMITED TO NEW PROPERTY...

Rents will rise by an average 10 per cent every year

7200 fewer houses and flats will be built annually

Australia's economic output will shrink by 1 per cent

175,000 fewer jobs will be created over the next decade

Canberra's tax take will shrink by \$1.65 billion a year

The impact would go well beyond any (Budget) saving ... to a multitude of unintended consequences.







The @BISShrapnel report shows Labor's policy on #negativegearing will be bad for those who own a home, rent a home & invest in a home 8:37 AM - 3 Mar 2016 · Canberra, Australian Capital Territory



0



A golden age for data driven policy analysis?

- Linked administrative data is a game changer
- Large proprietary data sets will also be significant
- Both allow us to answer new questions and to form more timely policy responses

There are still barriers in using data to its full potential

- Data is still expensive, slow to arrive and sometimes difficult to access
- More expertise is required within the public service
- Poor quality 'black box' data analysis in contentious policy debates

Unleashing the golden age

- Boosting data culture embrace need for data availability and timeliness not just for admin data
- Boosting capability within all agencies to use the data
- Framework for access to proprietary data sets
- Countering poor quality analysis

Boosting 'data culture' –improving accessibility of administrative data

- Public service rightly has checks and balances on access to linked administrative data sets BUT often delays are bureaucratic in nature
- Greater recognition of the cost of poor data timeliness and access on research productivity and appetite (and ultimately quality policy decision making)

More funding for DataLab to:

- Increase capacity
- Improve data storage (e.g. by using databases rather than csv files)
- Increase staffing and boost technology for output vetting to allow it to be timely

 Autoritia Government

 Backbridte Gommand

 Data Availability and Use

 No. 82, 31 March 2017

Despite recent statements in favour of greater openness, many areas of Australia's public sector continue to exhibit a reluctance to share or release data.

The entrenched culture of risk aversion...greatly inhibits data discovery, analysis and use

29



The *Data Availability and Transparency Act 2022* is a step in the right direction

From the Office of the National Data Commission:

The objectives of the Act are to:

- serve the public interest by promoting better availability of public sector data
- enable the sharing of public sector data consistently with the <u>Privacy Act</u> <u>1988</u> and use of appropriate security safeguards
- enhance integrity and transparency in sharing public sector data
- build confidence in the use of public sector data, and
- establish institutional arrangements for sharing public sector data.

This will allow more (accredited) government and university users of public sector data for:

- delivery of government services;
- informing government policy and programs; and
- research and development.

...the scheme will support a modern data-based society, driving innovation and stimulating economic growth.

DAT Act 2022, Explanatory Memorandum



Boosting data culture (cont) – improving access to public data



Publicly-available data should be easily accessible

- Well documented
- Consistent methodologies and naming of series over time
- Machine-readable, with APIs to easily access data
- Not (only) published on interactive dashboards without downloadable data

Anonymised microdata should be public where safe; or available within restricted environments when not

• Detailed aggregates should also be made available

Modelling code and data should be made public where safe; or just code where data are sensitive

 Governments should make their analyses clear and transparent



Adam Fennessey, Dean ANZOG

We've got to put our minds to more innovative use of de-identified data for quicker turnaround opportunities, to use data to better public policy without being bogged down in lots of argument and debate..... There are a lot of risks in data and Optus has shown us that, but how do we keep pushing the mindset back to how do we make better decisions for communities with data?

The good.....



HOME / ABOUT PARLIAMENT / PARLIAMENTARY DEPARTMENTS / PARLIAMENTARY BUDGET OFFICE / DATA PORTAL

Data portal

The PBO data portal includes historical fiscal data and datasets published as part of research publications. These have been consolidated below for ease of access and to support transparency around fiscal and budget policy issues.

If you find this data portal useful, have suggestions for improvement or additional data you would like to see in this portal, please provide your feedback to feedback@pbo.gov.au.

Historical fiscal data

Historical data published in the government's latest Budget update are available here in an accessible form. The data includes time-series not previously available in a single location.

Update	Data download	Date issued
2022-23 October Budget	🔳 (479 KB)	25 October 2022
2022-23 March Budget	🔳 (210 KB)	29 March 2022
2021-22 Mid-Year Economic and Fiscal Outlook	🔳 (189 KB)	24 February 2022

Regular PBO publications

Data for the PBO's regular annual publications (Beyond the budget (previously the medium-term fiscal projections), National fiscal outlook, Budget chart pack, and Budget and Mid-Year Economic and Fiscal Outlook (MYEFO) snapshots) are available below.

Beyond the budget: Fiscal outlook and projections

(known as the medium-term fiscal projections report until 2020-21)

Provides analysis and projections of key Commonwealth Government budget outcomes such as debt and the budget balance, and detailed projections of receipts and payments over the medium-term period.

From 2020-22, the report was expanded to include an assessment of fiscal sustainability over the next forty years.

Data for each of the editions are available below.

Publication	Data download	Date issued
Beyond the budget 2021-22: Fiscal outlook and projections	(<u>521 KB</u>)	21 September 2021
2020–21 Medium-term fiscal projections	🗐 (94 KB)	10 December 2020
2019–20 Medium-term fiscal projections	🔳 (59 KB)	26 September 2019
2018–19 Medium-term fiscal projections	🔳 (53 KB)	6 September 2018
2017–18 Medium-term fiscal projections	🔳 (694 KB)	5 July 2017
2015–16 Medium-term fiscal projections	🔳 (18 KB)	24 June 2015

Parliamentary Budget Office
About the PBO
What's new
COVID-19 publications
Frequently asked questions
Costings and budget information
Publications
PDI Calculator
PBO conference
Guides to the budget and PBO work
Online budget glossary
Data portal
General elections
Guidance for Senators and Members
Guidance for Commonwealth Agencies
Careers
Contact Us



All data together in a single place

Clearly labelled spreadsheets

Clean data

32

The less good





There are thousands of lines of data in the Australian Electoral Commission disclosures, and the information is hard to sort and categorise.

- Income from fundraising events is grouped with 'other receipts', which it impossible to distinguish from all sorts of other income, such as bank interest
- The same donor can appear multiple times under various names (e.g. 'SDA', 'Shop, Distributive and Allied Employees Association', spelling mistakes, various branches, missing data etc.)

Many initiatives underway to boost public service expertise – need to make the APS a great place for data analysts to work





Significant number of worthwhile initiatives already underway:

- APS Data Profession Stream
- Data Capability Framework
- Data Fellowships
- Digital Apprenticeships

Also need to make the APS a great place for data analysts to work

- Given them the tools
- Give them good management and hands on learning
- Embrace better processes like Reproduceable Analytical Pipelines
- Opportunities to publish research

Building in-house consultancy capability also promising

Should emphasise data and modelling capability

Agencies like PBO and Productivity Commission also important role

 Independences to ask right questions and the resources to do high quality modelling



- Some private organisations provide trusted users within government/researchers access to their data sets – 'data philanthropy'
- This is welcome but risk is government too reliant on goodwill from private businesses
- Burning question what circumstances does data become so important that government should be able to regulate/compel access for public interest uses?



- Some state governments publish guidelines for CBA assessments, would be good to have similar guidelines or 'minimum practice standards' nationally
 - Transparency on key assumptions (e.g. elasticities), clearly detailed methodology etc
 - Modelling code and data should be made public where safe; or just code where data are sensitive
 - Accountability (publish names of responsible partners)
- Government could require that modelling for government meets these minimum standards
- For third party modelling, standards help provide guardrails for the public debate – red flag if modelling does not meet minimum standards





High-quality public data analysis also helps counteract misleading statements/modelling



- Alternative approach to counteracting bad modelling is to have more dispassionate facts/analysis on the public record to counter spurious claims
- PBO is uniquely placed to play an 'honest broker' role in public debates within its remit
- Fore example, distributional analysis in franking credit rebates debate

Request for budget analysis

Dividend imputation credit refunds – further information				
Person/party requesting the analysis:	The Hon Matt Thistlethwaite MP, Member for Kingsford Smith			
Date analysis completed:	19 November 2018			
Expiry date of the analysis:	Release of the next economic and fiscal outlook report.			
Status at time of request:	Confidential Not confidential			

Summary of request:

The request asked a series of questions in relation to the Parliamentary Budget Office's (PBO) publicly released response to Senator David Leyonhjelm's request, *Dividend imputation credit refunds* (PBO response of 4 May 2018, published on the PBO website 8 May 2018, PBO reference PR18/00145).

Will Labor's dividend imputation policy overwhelmingly affect the low paid?





In a speech to the Alliance for a Fairer Retirement System, Assistant Treasurer Stuart Pohert said the plan would bit some of the lowest paid Australians

Tax relief for rich superannuation holders costs budget \$6b a year: analysis

Heath Aston and political correspondent Updated April 8, 2015 — 3.42pm, first published at 12.15am

Save A Share A A

Tax breaks enjoyed by mainly wealthy superannuation account holders are costing the federal budget \$6 billion a year in lost revenue, independent analysis shows.

The true extent of the damage to tax receipts by a system which allows tax-free earnings for those over 60 is revealed in research by the Parliamentary Budget Office. Providing solutions to Australia's most pressing policy issues





But out-of-pocket costs for out-of-hospital medical services have increased by 50 per cent on average over the past 10 years



Average out-of-pocket cost per type of service in 2010-11 and 2020-21, inflation adjusted



Many Australians – especially those with chronic health conditions – miss getting needed care or prescriptions due to cost



The estimated number of adult Australians (based on survey findings) who said they had missed or delayed a prescription or skipped or delayed care because of cost, 2020-21



Source: Grattan analysis of ABS Patient Experiences Survey, 2020-21. Sample: 28,386

Younger people – particularly younger women – are much more likely to skip care because of cost



Proportion of people who said they had missed a prescription or skipped or delayed needed care in the because of cost, 2020-21



Source: Grattan analysis of Patient Experiences Survey, 2020-21. Sample: 28,386

Our MADIP dataset included the following datasets:

- Core module
- Geography
- Medicare Benefits Schedule (MBS), 2011 to 2019
- Pharmaceutical Benefits Scheme (PBS), 2011 to 2019
- Census of Population and Housing 2016
- Personal Income Tax
- Centrelink Administrative Data



ID	Age	Sex	Postcode	Income	Welfare	Medicare spending	Medicare services	Medicines purchased	Cost of medicines	etc
XX	60	F	3055	\$120k	0	\$350	GP consult	Class A	\$75	