TRIALS AREN'T EXPENSIVE, DATA IS EXPENSIVE

Professor Nicholas Biddle

School of Politics and International Relations Research School of Social Sciences



The evaluation problem

 Impact evaluation is essentially a problem of missing data – cannot observe outcomes in absence of program. A person cannot be treated and not be treated at the same time!



Only solutions:

- Compare outcomes between groups of different people at the same time.
- Compare outcomes for the same group of people at different times ("before-and-after").



The evaluation solution?





Evaluation barriers

Strongly disagree Disa	ag ree	Agre	e 🗆 S	strongly	agree						
	3.8										
Randomly choosing whether some people get a policy intervention and others do not is unfair			35.8				49.	.2		11	2
	6.9										
Controlled experiments or trials are too expensive as ways of designing and testing social policies	0.5				61.8				28.3	3	3.0
	1.0										
It is important to rigorously evaluate social policies before applying them to everyone in the population	6.8	3			62.3				29	0.9	
	1.1										
Evaluations of social policies should be done independent of governments		16.9				64.1				17.9	
	1.4										
Governments cannot be trusted to evaluate their own policies		22.4				56.9				19.3	
	2.7										
Even if an intervention hasn't been trialled, it is unethical to not give someone a policy intervention if the government can afford it and it thinks it will work	3.7		38.1					52.1			6.1
Most new social policy interventions improve the outcomes of those who receive them	3.0		37.9					56.5			2.7
	0.0	10.0	20.0	30.0	40.0	50.0	60.0	70.0	80.0	90.0	100.0

Biddle, N., Gray, M. and Hiscox, M., 2023. Public support for Randomised Controlled Trials and nudge interventions in Australian social policy. https://csrm.cass.anu.edu.au/research/publications/public-support-randomised-controlled-trials-and-nudge-interventions-australian



The admin data environment I – PLIDA





5

The admin data challenge

- Current use of PLIDA has been mostly descriptive, or at best quasiexperimental
 - 230 active PLIDA projects on ABS list, only 1 has the word 'random' in the description (Employment Services Trials (2023))
- Harron et. al (2017) identified a number of general challenges in using admin data for research
- Fahridin et al. (2024) undertook a systematic review to 'describe how and why administrative data have been used in Australian randomised controlled trial conduct and analyses'
 - 36 RCTs identified, all within health. Majority used State/Territory data
 - Advantages minimises participant recall bias, reduces loss to follow-up, pre-trial analysis particularly for cluster RCTs
 - Disadvantages missing data, cross-jurisdiction mobility, construct validity, biases in consent to data linkage, time delays

[•] Fahridin, S., Agarwal, N., Bracken, K., Law, S. and Morton, R.L., 2024. The use of linked administrative data in Australian randomised controlled trials: A scoping review. Clinical Trials, p.17407745231225618.

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M.L. and Goldstein, H., 2017. Challenges in administrative data linkage for research. Big data & society, 4(2), p.2053951717745678.

The role of surveys in RCTs

- Classic Undertake surveys of treatment and control groups
- Methodological The effect of different survey approaches on survey outcomes
- Population based survey experiments (Mutz 2011)
- 'Control' group where data is only observed for treatment (or control group has missing data biases)

8

Survey challenges I – Ethics and distress (set-up) 9

- At start of COVID-19, there were concerns that asking surveys, including but not limited to COVID-19 experience would cause distress
 - ANU ran a simple RCT to test for impact
 - Sollis, K., Biddle, N., Edwards, B. and Herz, D., 2021. COVID-19 survey participation and wellbeing: a survey experiment. Journal of Empirical Research on Human Research Ethics, 16(3), pp.179-187.
- Self-reported questions replicated Gibbs et al. (2018) to measure the impacts of bushfires in Victoria. Asked at the end of the survey:
 - "How distressing did you find this survey?"
 - "We are interested to know whether you are glad that you participated in this survey?".
- Responses were on an 11-point Likert scale from 0 to 10, where 0 is "not at all", and 10 is "extremely".
- Life satisfaction
 - "Overall, how satisfied are you with life as a whole these days?".
- Respondents can answer on a scale from 0 to 10, 0 being "not at all satisfied", and 10 being "completely satisfied".
- 80% (n=2,534) asked as third question (after satisfaction with direction of country and voting intentions);
- 20% (n=621) asked as third-last question (before participant experience questions)
 - No statistically significant predictors of random assignment (balanced treatment and control groups)



Survey challenges I – Ethics and distress (results)

- The average level of life satisfaction for those who completed the question at the start of the survey was 6.50.
- For those who completed the life satisfaction questions at the end of the survey, the average was 6.63 (p-value = 0.260)
- The small population-level improvement in subjective wellbeing from completing the survey is being driven by particular population groups.
 - A significant positive difference was observed for females, those aged 55-74, and those who responded early in the survey period.
- Some population groups showed worsened reported wellbeing post-survey completion.
 - Individuals who reported to have probable serious mental illness; those who were living in financial insecurity; and those in the lowest decile of income



Survey challenges II – Response rates 11

- Response rates have been declining for most surveys that would be used as part of an RCT
 - Official statistics, e.g. National Health Survey:
 - 2017-18 = 76.1%, 2020/21 = 34.6%, 2022 = 56.7%
 - Quasi-official statistics, e.g. HILDA:
 - 2002 = 81.4%, 2020 = 60.2%
 - Academic surveys, e.g. ANUpoll:
 - 2008 = 31.8%, 2016 = 21.5%, 2019 = 8.6%, 2024 = 3.7%
- Range of trials and experiments to test for:
 - Boosting response rates through data analytics -<u>https://www.abs.gov.au/statistics/research/raising-survey-</u> <u>response-rates-using-machine-learning-predict-gold-providers</u>
 - Comparison of accuracy of results from different survey modes and recruitment methods - <u>https://srcentre.com.au/wp-</u> <u>content/uploads/2023/12/ACSSM-Analytical-Report-20231005.pdf</u>
 - Longitudinal follow-up ANU study showed that random assignment of social norms (x% of people had completed Wave 2) increased 24 hours completion rate from 42.0 to 48.7 per cent

Survey challenges III – Data linkage consent

- Edwards and Biddle (2021) conducted survey experiment to test impact of data linkage consent on survey outcomes and predictors of consent
- RQ1. Do requests for data linkage consent affect response rates in subsequent waves?
- RQ2. Do consent rates depend on (a) survey mode (online versus other); (b) types of data linkage requested; and (c) the length of the request?
- RQ3. Is understanding of the data linkage process effected by the type or length of data linkage request? Does understanding vary by other survey and demographic characteristics?
- RQ4. Are perceptions of the risk of data linkage effected by the type or length of data linkage request? Does risk vary by other survey and demographic characteristics?



Edwards, B. and Biddle, N., 2021. Consent to data linkage: Experimental evidence from an online panel. Advances in longitudinal survey methodology, pp.181-203.