

About the research

Data linkage and statistical matching: options for the Longitudinal Surveys of Australian Youth

Sinan Gemici and Nhi Nguyen, NCVER

Recent evaluations of the Longitudinal Surveys of Australian Youth (LSAY) have recommended investigating the potential for combining LSAY data with external data sources as a way to improve the breadth of information in the survey, but without adding respondent burden. Against this backdrop, the purpose of this discussion paper is to investigate the potential for linking data from existing administrative collections to LSAY and to explore the viability of combining data from LSAY and the Longitudinal Study of Australian Children (LSAC).

Key messages

- Linking administrative data from the education, training and health sectors would greatly enhance the ability to explore key drivers of young people's transition outcomes in LSAY without increasing respondent burden.
- The potential benefits are particularly appealing in topic areas that are currently quite limited in LSAY, such as health information, childhood development and early education outcomes. This makes linking the National Assessment Program – Literacy and Numeracy (NAPLAN) and Medicare data to LSAY the most valuable initial option.
- In a further stage, linking data from the Department of Human Services (Centrelink), the Australian Census, and national education and training statistics to LSAY could provide an evidence base for generating insights into the intergenerational impact of disadvantage.
- Although a statistical match between the Longitudinal Study of Australian Children and LSAY is at first sight appealing, given the complementary nature of these two flagship surveys, a closer look reveals a number of methodological obstacles. Research findings from such an amalgamated dataset of 'synthetic' individuals would lack the necessary robustness to inform evidence-based policy.

Overall, strong consideration should be given to concrete plans for linking administrative collections to LSAY, beginning with NAPLAN and Medicare data.

Tom Karmel
Managing Director, NCVER