As an Australian PhD student in economics, I have some experience working with education data and attempting to evaluate the impacts of policies and other environmental changes on student outcomes. There are several issues that I see as posing major problems for research in this area in Australia. In some cases such as data collection and privacy protection, you capture the problems and potential solutions very well in your draft report. In other areas I have some suggestions, which you may choose to investigate further and incorporate in the final version.

Privacy Protections and Data Linkage

The issue of privacy protections was discussed thoroughly in the report and the recommendations address the existing problems fairly effectively. However, I would add to Draft Recommendation 5.1 that the importance of collecting consent for research purposes extends beyond agencies involved in education data collection to any organization that collects information that could be used to measure future outcomes. Important examples of this are the Australian Tax Office (ATO), universities and health departments. The ability to link these data to educational interventions would allow researchers to evaluate the effects of interventions on outcomes that are far more meaningful than test scores, such as wages, university attendance and health.

Data linkage of this kind is essential to confirm that policies such as class size reductions (Chetty, 2011) or value-added-based teacher accountability measures (Chetty et al, 2013) have impacts on later-life outcomes despite the effect on test scores often fading out in subsequent years. Even more ambitious projects are possible in countries such as Norway and Sweden, which are in some ways a benchmark for data availability. For example, the linking of sensitive data allowed Kirkeboen et al (2016) to accurately measure the returns to choosing specific university courses and higher education institutions.[1] Such a project is impossible in Australia despite its clear value and an institutional structure that should allow a similar research design.

A second suggestion is that clear procedures be put in place for when sensitive data from multiple institutions across different domains (or jurisdictions) need to be linked. For example, individual students may need to be linked to incomes reported to the ATO. This does not seem to be possible with the data linkage nodes currently in existence. One possibility is that one institution could obtain access to data from the other and perform the linkage and de-identification in-house, but this relies on changes in privacy protections and high levels of cooperation from institutions that have little incentive to provide such a service. A potentially superior solution in my view would be to have a dedicated federal office in charge of linking datasets and releasing de-identified data to researchers.

Data Quality, Surveys and Censuses

The report suggests that surveys and samples could be used instead of censuses to increase the breadth of data at the cost of depth. This may in some cases be a sensible option. However, it should be stressed that surveys need to be representative. When surveys are used instead of censuses, non-response and non-random sampling can prevent researchers from gaining an understanding of the causal effect of the policy they are trying to evaluate regardless of the size of datasets and the variables available.

[1] Norway and Sweden are in many ways benchmarks in data availability. While strong restrictions are in place to protect privacy, access to an extensive range of linked individual-level data has facilitated unprecendented research that no doubt provides invaluable information to policymakers in those countries.

Randomized Control Trials

The most striking phenomenon I have noticed about education research in Australia is that randomized controlled trials (RCTs) are rarely used. The report raised the issue that these trials are sometimes rejected on the grounds that one group receives treatment and another does not. However, this is also true of "pilot" projects, which are identical in this regard but are much less useful for evaluation. Oddly enough, such pilot studies are typically not criticized on these grounds.

A closely related point is that government initiatives, including but not limited to those in the area of education, could also be tested using RCTs. This would make it possible to rigorously evaluate policies such as the many components of the Northern Territory Intervention, which is unfortunately exceedingly difficult due to the way in which the Intervention was implemented. Perhaps closer cooperation with academic researchers would help government agencies implement policies in ways that allow them to be effectively evaluated.

References

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011), "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR", *Quarterly Journal of Economics*, 126(4), 1593–1660.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood", *American Economic Review*, 104(9), 2633-2679.

Kirkeboen, L., Leuven, E. and Mogstad, M. (2016), "Field of Study, Earnings, and Self-Selection", *Quarterly Journal of Economics*, Advance Access.

Craig Ashley

8 September 2016