DISCUSSION PAPER

NCVER

Enhancements to the Longitudinal Surveys of Australian Youth

NATIONAL CENTRE FOR VOCATIONAL EDUCATION RESEARCH

Longitudinal Surveys of Australian Youth

Longitudinal Surveys of Australian Youth

Publisher's note

This paper was completed in December 2013 and was current at the time of writing. It was prepared by the National Centre for Vocational Education Research (NCVER) and commissioned by the Australian Government Department of Education to help inform the Longitudinal Surveys of Australian Youth (LSAY) review conducted in 2013-14.

The remit of the LSAY review was to consider how the LSAY survey has been used, whether it provided value for money and how it could be improved and made more useful for the evolving policy environment.

Additional information relating to this research is available in *A review of the Longitudinal Surveys of Australian Youth*. It can be accessed from the LSAY website < http://www.lsay.edu.au/publications/2844.html>

This document should be attributed as NCVER 2016, *Enhancements to the Longitudinal Surveys of Australian Youth*, NCVER, Adelaide.

This work has been produced by NCVER through the Longitudinal Surveys of Australian Youth (LSAY) Program, on behalf of the Australian Government and state and territory governments, with funding provided through the Australian Department of Education and Training.

The views and opinions expressed in this document are those of the author and do not necessarily reflect the views of the Australian Government or state and territory governments.

COVER IMAGE: GETTY IMAGES/iStock

Published by NCVER, ABN 87 007 967 311

ISBN 978 1 925173 38 3

TD/TNC 122.13

Level 11, 33 King William Street, Adelaide, SA 5000
PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

Phone +61 8 8230 8400 **Fax** +61 8 8212 3436
Email ncver@ncver.edu.au **Web** <http://www.ncver.edu.au> <http://www.lsay.edu.au>
Follow us: <http://twitter.com/ncver> <http://www.linkedin.com/company/ncver>

About the research

Enhancements to the Longitudinal Surveys of Australian Youth

The Commonwealth Department of Education commissioned the National Centre for Vocational Education Research (NCVER) to assess the value and implications of eight enhancements to the Longitudinal Surveys of Australian Youth (LSAY).

The objective of LSAY is to track young Australians as they move from school into further study, work and other destinations to provide a meaningful dataset through which to understand youth transitions. Enhancements to LSAY are considered in this paper in the context of continuing to enable researchers to track young people over time and examine relationships between the variables that impact youth transitions.

Key messages

- All enhancements are interrelated, with a change to one aspect of LSAY affecting other aspects of LSAY.
- There are options to alter the sample design (including the frequency of starting new cohorts and changing the sample size) within the current survey design, whereby the sample is selected from school students who participate in the Programme for International Student Assessment (PISA).
- Addressing attrition from the first wave (PISA) to the second wave will improve the value of the LSAY dataset and, in particular, improve the ability to analyse sub-populations.
- In order to make a case for gathering any new information, a relationship between the additional measures and the success of youth transitions needs to be demonstrated. There is support for improvements to information in areas such as outcomes beyond age 25, wellbeing and parental background.
  - Transitions are taking longer, providing support for extending the survey beyond age 25.
  - Higher levels of wellbeing are associated with more successful transitions to adulthood, providing support for improving the breadth of wellbeing information.
  - Variation between students on educational outcomes is related to family background but the quality of parental background information in LSAY is currently questionable because it is missing or inaccurate.
- Collecting or improving information on outcomes beyond age 25, health and wellbeing and parental background will incur costs and the accuracy of the information will be dependent on response rates. Various strategies can be undertaken to minimise costs and encourage participation.
- Linking LSAY with administrative collections allows for the inclusion of information from other time dimensions and the improvement of information in areas identified as weak, including health and wellbeing and parental background.

Contents

Tables and figures

Tables

Figures

Executive summary

The Commonwealth Department of Education[1] commissioned NCVER to examine and report on areas in which LSAY could be enhanced. The findings from this report will be considered by the department in the 2013 review of LSAY. The remit of the department's review of LSAY is to consider how the survey has been used, whether it has provided value for money and how it could be improved and made more useful for the evolving policy environment. (See the terms of reference in appendix A.)

Background information

The LSAY program commenced in 1995 and was based on two other annual surveys: the Australian Youth Survey (AYS; 1989—97) and the Youth in Transition survey (YIT; 1978—1996). Survey participants (known as a 'cohort') enter the study when they turn 15 years, or as was the case in earlier studies, when they were in Year 9. Individuals are contacted once a year for ten years. Since 2003, the initial survey wave has been integrated with the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA). Over 10 000 students start out in each cohort.

A longitudinal survey obtains information from the same respondents at multiple points in time (Bureau of Labour Statistics 2013). This enables researchers to:

- account for the unobserved differences amongst cohorts and to investigate and make inferences about the relationships between the variables of interest;
- isolate the influence of policies and practice from confounding influences such as social background and context;
- track patterns of development and change over time;
- identify sequences and pathways; and
- identify critical periods in human development for exposures and risks and inform the timing of preventive measures (Logie, Hogan & Puet 2004; Bureau of Labour Statistics 2013; Howieson, Croxford & Howat 2008).

The purpose of LSAY is to track young people as they move from school into further study, work and other destinations to provide a meaningful dataset through which to understand youth transitions. It is in this context that the advantages and disadvantages of possible enhancements to LSAY are considered.

In 2009, Nguyen et al. (2010) assessed the fitness for purpose of LSAY in generating a data source that could be used to understand youth transitions. This 'stocktake' review resulted in five broad recommendations to improve the usefulness of LSAY, ranging from reviewing the survey content to improving survey attrition. This discussion paper builds on the stocktake report by Nguyen et al. by providing an assessment of the value and implications of the recommended enhancements within the remit: that the main purpose of LSAY is to provide an information source to enable young people to be tracked from school into further education and work.

This paper assesses the followings enhancements:

- reconsidering the frequency for starting new LSAY cohorts

[1] Formerly known as the Department of Education, Employment and Workplace Relations (DEEWR).

- evaluating and changing the sample size
- adopting measures to reduce attrition
- extending the age to which LSAY cohorts are followed to beyond 25 years
- introducing a parent questionnaire to collect more comprehensive background information on respondents
- reviewing the survey questionnaires to improve data collection on health and wellbeing, resilience and adaptability
- linking to other educational and administrative datasets such as the NAPLAN
- introducing supplementary topical surveys, interviews, focus groups or other means of enhancing the usefulness of LSAY to policy-makers.

The methodology involved assessing the advantages and disadvantages as well as the feasibility of each enhancement through data analysis, technical knowledge and a literature review. For the purpose of this paper, the changes that come with substantial costs have been noted, although we have not allowed costs to constrain our thinking.

Findings

As noted, the purpose of LSAY is to track young Australians as they move from school into further study, work and other destinations. LSAY is not a whole of youth study, and, as such, the content of the data and questionnaires should be restricted to the elements that measure and impact on education and employment.

In their earlier review, Nguyen et al. (2010) noted that this ability to track the population of interest and examine the relationships between variables that influence transitions makes LSAY a valuable dataset for researchers and policy-makers. Our assessment confirms that the survey design produces robust estimates of the population at the national, state and territory and school sector levels. There are, as always, some limitations. The options for addressing these limitations are considered in this paper but in the context of continuing to enable researchers to track young people over time and examine relationships between the variables that impact youth transitions.

The body of the report provides an assessment of each enhancement. It should be noted that all enhancements are interrelated, with a change to one aspect of LSAY affecting other aspects of LSAY. For example, extending the age to which cohorts are followed must be weighed against the impacts of attrition over a longer timeframe. It should also be noted that a major change to the LSAY survey design may involve a prioritisation of other improvements, have cost implications and require strategies to reduce respondent burden.

Sample design and maintenance

LSAY survey participants are currently selected from school students who participate in PISA. The options for changing the sample design (including the frequency of starting new cohorts and changing the sample size) and addressing attrition are more limited within the PISA arrangement.

On the other hand, the PISA sample design does offer a number of options for altering the sample size and reducing the frequency of introduction of new cohorts, with different PISA options available to the participating countries. For example, there is scope to extend the time between cohorts, to every six

years, which would free up funds to enhance the survey in other areas. While this would not detract from the key purpose of LSAY, a disadvantage would be a reduced capacity to examine the effects of economic downturns and other potentially influential national events that occur between cohorts. There is also scope to alter the sample size. Australia chooses to sample above the PISA minimum requirement (5000 students), resulting in a LSAY sample size of approximately 14 000 students. Although the current sample design works well in providing reliable estimates, there is scope to reduce its size for the larger states and increase the sample size for the smaller states, which would improve analyses at the jurisdictional level. If the LSAY—PISA link remains, a review of the sampling options within the PISA framework could be considered.

If LSAY is separated from PISA, there will be more control over sampling and the adoption of methods to reduce attrition. A key benefit would be the opportunity to change the sample design to make the existing sample more efficient. This could include improving the ability to further stratify schools and boosting the number of LSAY participants from equity groups. It is noted from the research that oversampling sub-populations could reduce the representation of the overall population of Australian youth. This potential consequence needs to be weighed against the main purpose of LSAY. Best practice also suggests it is preferable to examine specific sub-populations, such as the Indigenous population, via specialised surveys.

The key obstacles to breaking the LSAY—PISA link include costs, recruitment challenges, and information gaps on schools, students' background and academic performance. The information gaps would need to be addressed as a matter of priority through additional survey questions and testing. Of particular importance is the requirement for a reliable measure of academic performance, as research consistently demonstrates that literacy and numeracy are strong predictors of education and labour market outcomes. The most cost-efficient option for collecting information on academic performance is to obtain data from the National Assessment Program — Literacy and Numeracy (NAPLAN) through data linkage or to link the sampling of LSAY participants to NAPLAN testing. One of the first steps in assessing whether to maintain the PISA link would be to explore the accuracy of data linkage between NAPLAN data and LSAY and the reliability of NAPLAN scores as a predictor of youth transitions.[2]

Addressing attrition is fundamental to improving the value of the LSAY dataset and is linked to several of the enhancements considered in this paper. LSAY currently suffers from a loss of sample members, which is a common problem in longitudinal surveys. The most substantial sample loss occurs from the first wave (PISA) to the second wave, and represents a sample loss of between 20 and 40%. Field reports indicate that as much as 25% of a cohort can be lost because incorrect contact details are provided by students when sitting PISA. Addressing attrition from the first to the second wave is the area of most priority. This could be addressed by offering financial compensation to schools to provide accurate contact details, interviewing participants in the first wave or exploring options to provide incentives at this time.

The use of incentives is a method adopted by other comparable longitudinal surveys to encourage participation. The cost of providing incentives will be high if incentives are used for all LSAY participants (approximately 14 000 per cohort on commencement of the survey). The costs could be offset by reducing the sample size to a level where reliable and accurate estimates would still be produced — at least at the Australian and state and territory levels. Alternatively, incentives could be targeted to sub-populations, such as those suffering from the highest rate of attrition, noting that this raises equity issues.

Other strategies aimed at reducing attrition in all waves include establishing a strategic communication plan targeted to young people and their parents and rebuilding the sample or re-sampling sub-populations

[2] The findings from a forthcoming project by NCVER may provide insights into the correlations and differences between NAPLAN and PISA scores.

with large attrition. Furthermore, the implementation of some of these strategies has wider benefits. For example, the introduction of an information pack for participants and their parents has the capacity to promote the value of the survey and provides a means to administer parental surveys and letters of consent for data linkage and participation in ad-hoc surveys and studies.

Other enhancements

The remaining enhancements considered in this paper are related to improvements to the information in areas such as outcomes beyond age 25, health and wellbeing and parental background. The evidence suggests that improved information in most of these areas is warranted and would result in a dataset that would further enhance research on youth transitions. Research confirms that transitions are taking longer, supporting the value of extending the survey beyond age 25. There is evidence that higher levels of wellbeing are associated with more successful transitions. There is less evidence of a relationship between health measures, such as dietary intake and physical activity, and successful youth transitions. Finally, research shows that approximately 80% of the variation in educational outcomes is linked to individual and family background factors. Data on parental background are collected in the first wave of LSAY. The quality of this information is questionable because it is reported by the young people themselves, whom may not be fully aware of their primary care giver's qualifications or occupation.

There are several options for collecting or enhancing information on outcomes beyond age 25, health and wellbeing and parental background. As noted, extending the survey beyond age 25 is likely to impact on attrition and will increase survey costs. A possible approach to minimising costs would be to survey people on a biennial basis after age 20. This would result in an increase of only one extra survey wave, but extra costs would be incurred through increased sample maintenance. Information gaps on health and wellbeing could be addressed by the inclusion of a health and wellbeing module in the questionnaire, either through a personal interview or via the current computer-assisted telephone interviewing (CATI) mode. Other less costly approaches include asking a sub-sample of participants to participate in a separate study or investigating data linkage. This approach is particularly recommended for options to improve information on health because there is less evidence that diet and activity is related to the success of youth transitions. Response rates, testing burden and costs are the main barriers to the adoption of a parental questionnaire. A marketing plan and communication strategy may boost response rates, and costs could be reduced by exploring data-linkage options.

Data linkage could enable the information areas identified as weak, including health and wellbeing and parental background, to be improved by linking to existing administrative collections, such as Medicare and Centrelink. The key benefits to enhancing information through this approach are reduced costs and no impact on respondent burden. Privacy concerns and legal issues can be minimised by ensuring that consent is obtained and the use of an official data-integration authority. The conclusion is that it is technically possible to link LSAY with other datasets, which suggests many future possibilities for enhancing LSAY.

Summary

LSAY provides robust estimates of Australian youth and from this perspective is considered a valuable dataset for tracking young people over time. There are some limitations to the data (and longitudinal surveys in general). LSAY suffers from attrition, particularly from the first wave (PISA) to the second wave. It is limited in its ability to provide in-depth analyses of health and wellbeing, to capture accurate information on parental background and, due to high attrition, to provide reliable estimates of the Indigenous population below the national level.

The specific enhancements considered in this paper (for example, changes to the sample design, reducing attrition, linking to other datasets, and improving questionnaire content) are related, and all have merit, but the options and priority areas may vary depending on whether LSAY continues to be linked to PISA. Other key considerations to enhance the value and usage of LSAY include addressing attrition and linking to other datasets. The cost of the options will be an important issue, with some options bearing significant costs, without guarantee that the benefit to the dataset will be of the same magnitude. Priority areas can be identified by considering the findings from this paper and the wider review of LSAY being conducted in parallel to the preparation of this paper.

The remainder of the document considers each enhancement, in detail, in a separate chapter. Based on the literature, analyses of LSAY data and technical opinion, the following elements are considered for each chapter:

- advantages and disadvantages of the change;
- technical and/or practical feasibility;
- influence on analysis and reporting; and
- approximate costs.

The paper ends with concluding comments.

Dr Craig Fowler
Managing Director, NCVER

1. Frequency of starting LSAY cohorts

An important issue for consideration is the frequency with which to start new LSAY cohorts. LSAY is currently implemented as a continuation of the PISA study, which means that new LSAY cohorts are automatically tied to the triennial PISA testing cycle. Therefore, an important question is whether to continue the integration of LSAY with PISA or whether to separate LSAY from PISA. Both options offer distinct advantages, as outlined below.

Option 1: Maintaining the LSAY–PISA link

As noted, the beginning of new LSAY cohorts is currently linked to the three-yearly administration of the PISA study. As a consequence, there are currently few opportunities to alter the frequency of new participating cohorts. Nevertheless, it would possible to skip a PISA cycle and begin a new LSAY cohort every six rather than every three years. For instance, no new LSAY cohort was started with the 2012 administration of PISA. However, there are problems associated with skipping PISA cycles. The extended time between cohorts (that is, six years) may lead to gaps in capturing the effects of major social, cultural and macro-economic trends on young people's transition decisions. One prominent example is the Global Financial Crisis (GFC) in 2008, which resulted in a complete collapse of the labour market for young people (see Karmel 2013). Drastic and unforeseen events such as the GFC are likely have a powerful impact on young people's pathways, and the full extent of this impact is best measured via inter-cohort comparisons. Therefore, extending the time period between cohorts may result in a loss of important information, such as the length of time spent looking for work or changes in the uptake of government support payments.

It would also be possible to follow a single cohort. This would result in cost savings that could be redirected towards other improvements to the survey, but there are some obvious disadvantages to this approach. It is not possible to measure and compare cohort and period effects under this design. Logie, Hogan and Puet (2004) add a further disadvantage: single-cohort studies are slower to yield longitudinal results because the population must be followed through a number of waves in order to observe effects over time.

While the relationship between PISA and LSAY removes flexibility in terms of cohort frequency, the link offers a number of important advantages.

Academic performance measure

One benefit of using PISA as the first wave of every LSAY cohort is that it provides an objective measure of students' literacy and numeracy performance at age 15.[3] The ability to control for academic performance is crucial because it allows researchers to identify causal relationships between young people's pathways and their ultimate transition outcomes. Longitudinal surveys without academic performance measures are hampered by differences in young people's ability, making it very difficult to isolate causal effects.

[3] Prior to the connection with PISA, the Y95 and Y98 cohorts of LSAY used proprietary test questions to obtain information on literacy and numeracy. However, the literacy/numeracy component for these cohorts was considerably less robust than the assessment in PISA.

PISA literacy and numeracy scores are based on robust OECD methodology and are widely accepted as a high-quality measure of academic performance. Moreover, they allow comparisons between the academic performance of 15-year-olds across more than 70 countries. One repercussion of separating LSAY from PISA would be losing the capacity to carry out international comparisons, although this is more pertinent to performance on PISA scores than with respect to youth transitions. Another consequence would be having to find an equally well-established substitute measure of academic performance for LSAY. Obtaining such a measure would have cost and resource implications.

Sampling and data collection

PISA uses a tried and proven sampling strategy to obtain a nationally representative sample of 15-year-olds, with all the requisite processes for sampling, recruitment and data collection now firmly established. Separating LSAY from PISA would mean that sampling, recruitment and data collection procedures would have to be re-developed and re-implemented, which would have cost and resource implications.

Schools data

Another advantage of the current model of LSAY—PISA integration is the availability of detailed data on the over 350 schools whose students participate in the PISA study. School questionnaires are completed by their respective principals, and data are gathered on a wide range of relevant topics, including school structure and organisation, teacher and principal characteristics, resourcing, school climate etc. The ability to jointly analyse student and schools data is a highly desirable feature of the current LSAY—PISA link, in that schools have an important independent effect on young people's transition outcomes (see Lim, Gemici & Karmel 2013).

Another advantage of the PISA schools data file is that relevant variables, such as the presence of teacher shortages or the quality of educational materials, are statistically converted by PISA staff into continuous scores. This greatly facilitates data analysis and allows direct comparisons between schools.

Severing the link between LSAY and PISA would require the planning and implementing of a new LSAY schools questionnaire. It would further require dealing with technical issues such as complex statistical conversions for relevant variables. Finally, there would be a potential risk for increasing response burden in those cases where schools are selected independently for participation in both PISA and LSAY.

Option 2: Separating LSAY from PISA

As outlined above, there are strong arguments in favour of keeping the LSAY—PISA link intact. However, separating the two surveys offers opportunities for a fundamental overhaul of LSAY in a variety of areas.

Academic performance measure

A separation of LSAY and PISA would result in the loss of PISA's academic performance measures in reading, mathematics and science. The most adequate substitute for PISA is NAPLAN testing. NAPLAN was introduced in 2008 to assess all students in Years 3, 5, 7 and 9 in reading, writing, language conventions (spelling, grammar and punctuation) and numeracy. The integration of NAPLAN scores into the LSAY dataset could be achieved through a process known as *data linkage*. (Readers are referred to chapter 7 of this report for a detailed discussion of data linkage between LSAY and NAPLAN.)

The advantages of using NAPLAN scores as a measure of academic performance in LSAY are twofold.

- NAPLAN scores would provide access to young people's literacy and numeracy development from Years 3 to 9, allowing researchers to control for academic performance at earlier ages. This is an important advantage because early academic performance is among the most important predictors of later academic and labour market outcomes.
- Although PISA scores are widely accepted as a high-quality measure of academic performance at age 15, they assess literacy and numeracy in a very general sense. This is necessary because PISA is implemented in over 70 countries worldwide. The advantage of NAPLAN is that it is a purely Australian measure of academic performance, one which is specifically designed to align with national education priorities.

Sampling and data collection

PISA currently serves as the base year for each new LSAY cohort. In consequence, no direct control exists over sampling design, recruitment and data collection. Moreover, very little control exists over the design and content of the base year questionnaire, as these aspects are managed by international contractors who work under the auspices of the OECD secretariat and the PISA governing board.[4] Separating LSAY from PISA would offer various new possibilities with regard to sampling and data collection.

Using a school year-based sample

PISA targets young people at age 15. In consequence, the PISA sample is age-based rather than based on school year. The survey design for an age-based sample is considerably more complex because it needs to allow for the varied circumstances of young people, especially while at school (for example, 15-year-olds may be enrolled in any of Years 9, 10 or 11). This means that each year more questions in LSAY are now devoted to clarifying respondents' school year status and other individual circumstances. A model based on new cohorts' school year (for example, starting with Year 9 NAPLAN students) would reduce the survey's complexity and avoid potential complications during data analysis.

Oversampling students from all equity groups

PISA already oversamples Indigenous students to ensure that valid and reliable analysis of academic performance can be conducted. This oversampling of Indigenous students is important for LSAY because, over time, Indigenous respondents drop out of the survey at disproportionally high rates. Unfortunately, other equity groups that are also subject to higher attrition (for example, respondents from low-socio-economic status [SES] backgrounds, those with disabilities and those from refugee backgrounds) are not currently oversampled in PISA. By the time post-school transition takes place, many students from these equity groups have already dropped out of LSAY, leading to a loss of important information about their transition. This is problematic because the transition outcomes of equity groups are of great interest to policy and research. Separating LSAY from PISA would open up new possibilities to oversample all equity groups, resulting in more robust insights into how young people from equity groups are faring. See chapter 2 on sample sizes for further information on oversampling sub-populations.

Reducing initial sample size

Currently, the PISA sample consists of approximately 14 000 students. The size of the PISA sample is rather large compared with other well-known national and international panel studies, where considerably smaller initial sample sizes are used. (For examples of those studies refer to chapter 3 on attrition.)

[4] Some very limited opportunity for local adaptation in PISA is available through several national option questions.

Separation from PISA would enable LSAY to gain full control over its initial sample, possibly reducing sample size by up to one-third. The monetary saving resulting from a smaller initial sample could then be invested in sample maintenance, thus reducing the attrition problem inherent in LSAY. (Refer to chapter 3 on attrition.)

Gaining consent for data linkage

The prospect of linking LSAY to existing administrative databases is potentially one of the most promising and exciting undertakings in the years to come. (For a comprehensive description of data linkage with LSAY, please refer to chapter 7 of this report.) A major legal/ethical requirement for linking LSAY data is written consent from respondents or their parents/legal guardians in the base year. However, base year data are currently collected under the PISA contract, which contains no provisions for gaining written consent for future data linkage. Separating LSAY from PISA would transfer full control over the initial sample and data collection to LSAY administrators, enabling them to seek written consent for data linkage from the outset of every new cohort.

Conducting topical surveys, interviews and focus groups

Supplementary surveys, interviews and focus groups are often useful for illuminating specific youth transition issues. (See chapter 8 for further information on this approach.) Having full control of sampling and data collection would greatly facilitate the application of additional supplementary surveys and focus groups, thus further enhancing the usefulness of LSAY to policy-makers.

Cost considerations

Within the PISA sample design, it is possible to start a new cohort every three, six or nine years. Skipping a PISA cycle and beginning a new LSAY cohort every six years[5] rather than every three years would reduce data collection costs over a 15 year period by around a third, compared with a three-yearly cycle. For example, under the current sampling design, three cohorts (Y03, Y06 and Y09) will be surveyed for 11 waves between 2003 and 2018. Increasing the gap between cohorts from three to six years is equivalent to removing the data collection costs for all participants within the Y06 cohort. Based on the number of Y06 participants interviewed via CATI so far, it is estimated that the fieldwork for the Y06 cohort will involve over 50 000 CATI interviews from wave 2 to wave 11. [6] Increasing the gap between cohorts will also have cost savings in other areas, such as questionnaire development, dataset preparation and review, printing and postage of the sample maintenance product and the annual production of cohort reports and technical documents.

Separating LSAY from PISA offers unlimited options to commence a new cohort but this option has significant costs. As mentioned, there would be significant initial costs in areas such as sampling, participant recruitment, and questionnaire development and testing to capture information on schools. Separating LSAY from PISA would also increase fieldwork costs as data would need to be collected from school principals as well as students. Costs will be higher again if NAPLAN is assessed as being an

[5] Howieson, Croxford and Howat (2008) recommend a four to five year gap between cohorts to enable the analysis of social and policy change but without the costs associated with a more frequent recruitment of cohorts. The cost savings from implementing a nine-year gap between cohorts or following a single cohort are not covered in this section as there would be limited or no opportunity to capture cohort and period effects.

[6] CATI interviews are not conducted in wave 1 but combining the LSAY questionnaire with the PISA questionnaire in wave 1 incurs a cost of approximately $40 000.

inadequate substitute for the measurement of academic performance because a new measure would need to be obtained, tested and administered in wave 1.

Summary

A key question on cohort frequency is related to whether the current model of LSAY—PISA integration should be continued. The sections above have shown that both options have inherent advantages. The many benefits of the current model need to be balanced against keeping the frequency of new LSAY cohorts locked in with the triennial PISA cycle. By maintaining the PISA link, new cohorts could theoretically be initiated every three, six or nine years. Separating LSAY from PISA, on the other hand, would provide complete flexibility with cohort frequency, including annual and bi-annual new cohort options. However, it needs to be emphasised that a departure from the current model has cost and resource implications.

2. LSAY sample size

This chapter investigates LSAY sample size, and the implications of reducing or increasing the overall size of the sample. It should be noted that maintaining the LSAY—PISA link restricts scope for any change to the sample size or the sampling scheme.

The current sampling methodology is that of a two-stage stratified cluster design. Schools are sampled in the first stage (primary sampling units; PSUs) and students from each selected school are then sampled in the second stage. Schools are selected using probability proportional to size, so that large schools have a higher probability of selection than smaller schools.

In the current methodology, approximately 300 schools are chosen, with around 50 students from each school selected, which gives a total PISA—LSAY sample size of around 14 000 15-year-olds. The total number of 15-year-olds attending school is around 250 000 for each of the Y03, Y06 and Y09 cohorts. In 2012, the number of secondary schools from which schools could be selected was 2700.

The current multi-stage cluster design provides a convenient way by which to select the individual students, it more importantly allows the use of multi-level modelling to model the impacts that schools have on young people's outcomes. Further benefits of using the existing sample design include that it is consistent with the earlier LSAY cohorts, and that it has the capacity to utilise the schools to aid with recruitment and retention, to verify and clean student and parent information, and to provide school-level information (Collingwood et al. 2010).

Full details of the PISA sampling methodology can be found in the PISA technical manual (OECD 2009). It is worth noting that the sample size for the Australian PISA is twice as large as needed for PISA and was adopted to ensure adequate sample sizes for LSAY components.

The stratification of the LSAY sample occurs at the school level, whereby schools are divided into state/territory, school sector and regional (metropolitan, regional and rural/remote) strata. No further stratification is undertaken at the student level; however, all Indigenous 15-year-olds at a chosen school are selected to participate in PISA. That is, there is over-sampling of the Indigenous population to enable a meaningful national sample.

Information about schools is available from two major sources, the first being the Australian Bureau of Statistics schools collection (ABS 2012). The second is the Australian Curriculum, Assessment and Reporting Authority (ACARA), where data are primarily held for preparation of the MySchools website.

If the LSAY—PISA link were to be discontinued, it would be prudent to maintain the same sampling design for a redesigned LSAY, as the current cluster sampling scheme provides a convenient framework and there is more information about the schools that 15 year olds attend than about the population of 15 year olds themselves. Nevertheless, a redesigned sample frame would allow the addition of further information for defining different strata or groupings (such as low socioeconomic status) of schools and students when designing the sample frame.

The aim of PISA (and subsequently LSAY) is to provide broad population-level estimates at both the national (for PISA) and within states and sectors for LSAY. There is merit in LSAY having a sufficient sample size in key sub-populations of interest. For example, the oversampling of Indigenous students is an attempt to ensure reasonable sample sizes for this group of young people. As a general rule, however, boosting the sample sizes of sub-populations leads to inefficiencies in other areas, particularly other sub-

populations (Collingwood et al. 2010; Citro & Kalton 1993; Kalton 2009). Boosting sub-samples increases the overall sample size, and so population totals are estimated with a precision greater than needed.

In the following section, the impact of sample sizes on the precision of population estimates is examined. This is followed by a brief discussion on how it might be possible to design and boost samples for certain sub-populations.

Sampling and sample size

Cluster sampling is known to result in higher standard errors that would otherwise be obtained if simple random sampling had been undertaken (Dorefeev & Grant 2006). However, the benefits of using cluster sampling make it the most appropriate methodology for LSAY.

In this section, sample size calculations have been undertaken to determine the precision of population estimates. The sample sizes have been calculated individually for each LSAY strata of interest (state, school sector and the state by school sector combinations). These individual sample sizes are then summed to find national sample sizes. The calculations are based on the population sizes within each state and sector combination, using assumed design effects and historical response patterns.

In order to find appropriate population totals, the national schools collection, compiled by the ABS, has been used. This collection provides both the number of schools and the total number of 15-year-olds in each stratum in 2012. Should further decisions on sampling, sample size and population frames need to be made, then the full schools collection (and other data) should be obtained for analysis.

Further assumptions made in calculating the sample sizes included:

- a population proportion estimate of 50%
- a required margin of error of 5%; that is, the population proportion was to be estimated to be in the range of 45 and 55% with a 95% confidence level.
- design effect[7] = 2. A design effect of 1 equals the efficiency of a simple random sample.

Given that LSAY is a longitudinal survey, sample size calculations and their impact on the margin of error are presented assuming that the sample size is determined for wave 1. Sample size calculations were undertaken using the methodology outlined in Cochran (1977).

The results from the sample size calculations are presented as a series of figures. Each figure shows the obtained margin of error (precision) for a population estimate of 50% plotted against different sample sizes. The horizontal line at 5% reflects the targeted maximum margin of error of 5%. A graph is produced for each state, sector, state by sector combinations and for Australia as a whole. The results present the margins of errors for each sample size for wave 1 of the survey. However, information on later waves can be obtained simply by considering the response rate for each wave and then reading the respective margin of error. From the chapter on attrition (chapter 3), we know that wave 1 to wave 2 attrition is of the order of 20—40% and then wave-on-wave attrition is around 10—15%. It is estimated that by the last wave, around 30 — 35 % of the original cohort numbers remain.
Using figure 1, which represents the margin of error obtained for Australia across all sectors and for each of the school sectors separately, we see that for all sample sizes the margins of errors are below the recommended 5%. In order to determine the wave 10 (final) margin of errors, we would take the 15 000 or

[7] Design effect (DE) is the ratio of the variance of an estimate derived from a survey to the variance of an estimate of the same measure based on a simple random sample of the same size (Dorofeev & Grant 2006).

so and divide by 2, so that for the 7500 remaining in wave 10, we would obtain Australia-wide estimates with a margin of error of around 1.0. Thus, for all sample-size considerations, the Australia-wide estimates (both overall and within sectors) obtained produce reliable and accurate estimates. When considering each individual state and sector, it is observed that most of the estimates have precision at or below the recommended 5% for all waves. However, in the later waves, the precision of the smaller jurisdictions (Northern Territory, Tasmania and the Australian Capital Territory) would be greater than the recommended level, particularly for the individual school sectors of the Northern Territory.

From these results, there is scope to potentially reduce the sample sizes in some of the larger states without sacrificing precision; conversely, there is an argument for increasing the sample sizes of some of the smaller states to ensure that the impact of attrition on precision is reduced. That is, the current sample size works well in providing reliable population estimates, although there is some scope to redistribute the sizes of samples from the different state-by-sector combinations.

Figure 1 Margin of error for different sample sizes – Australia (assuming p = 0.5 and design effect = 2.0)

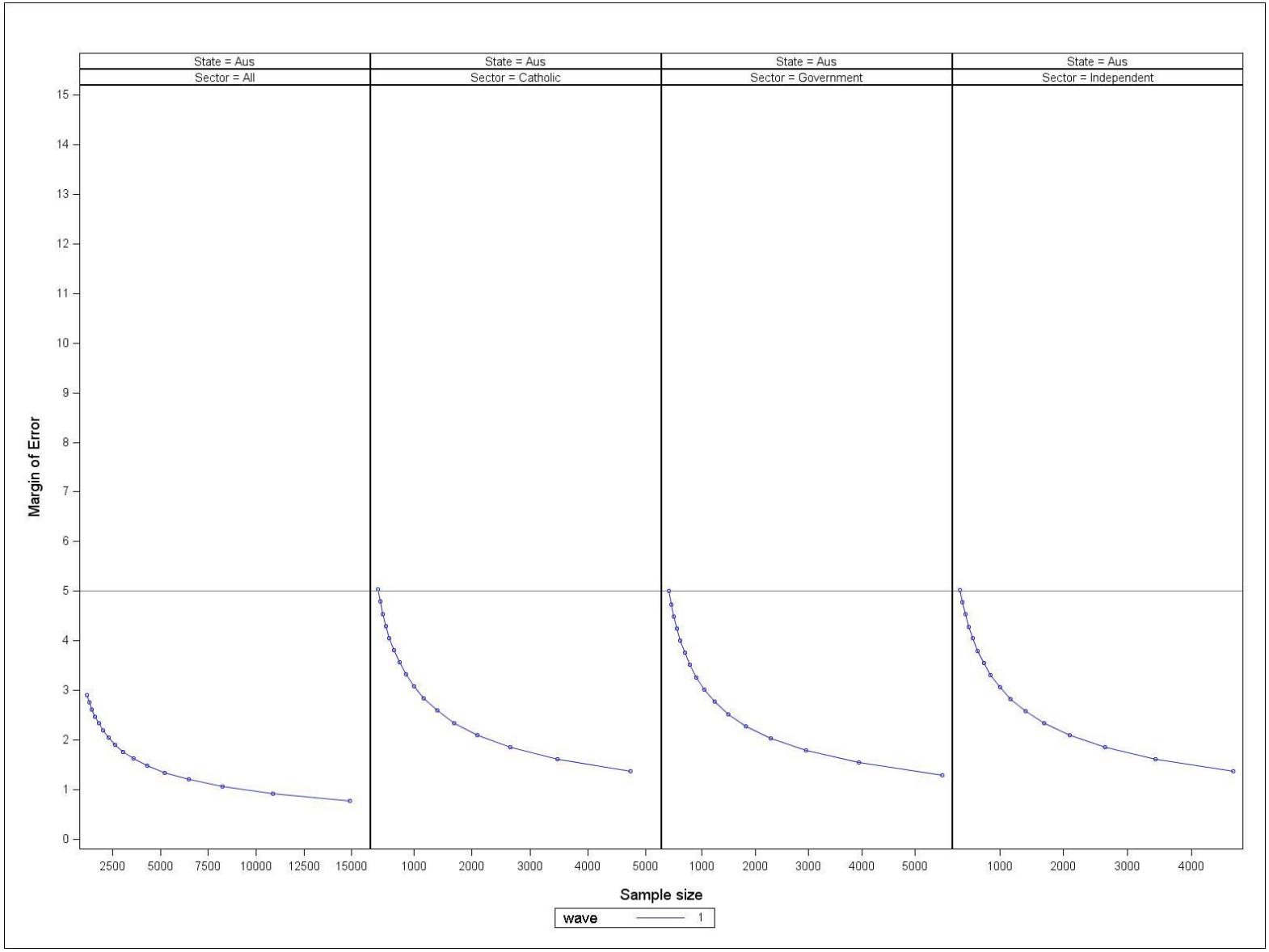


Figure 2 Margin of error for different sample sizes – ACT (assuming p = 0.5 and design effect = 2.0)

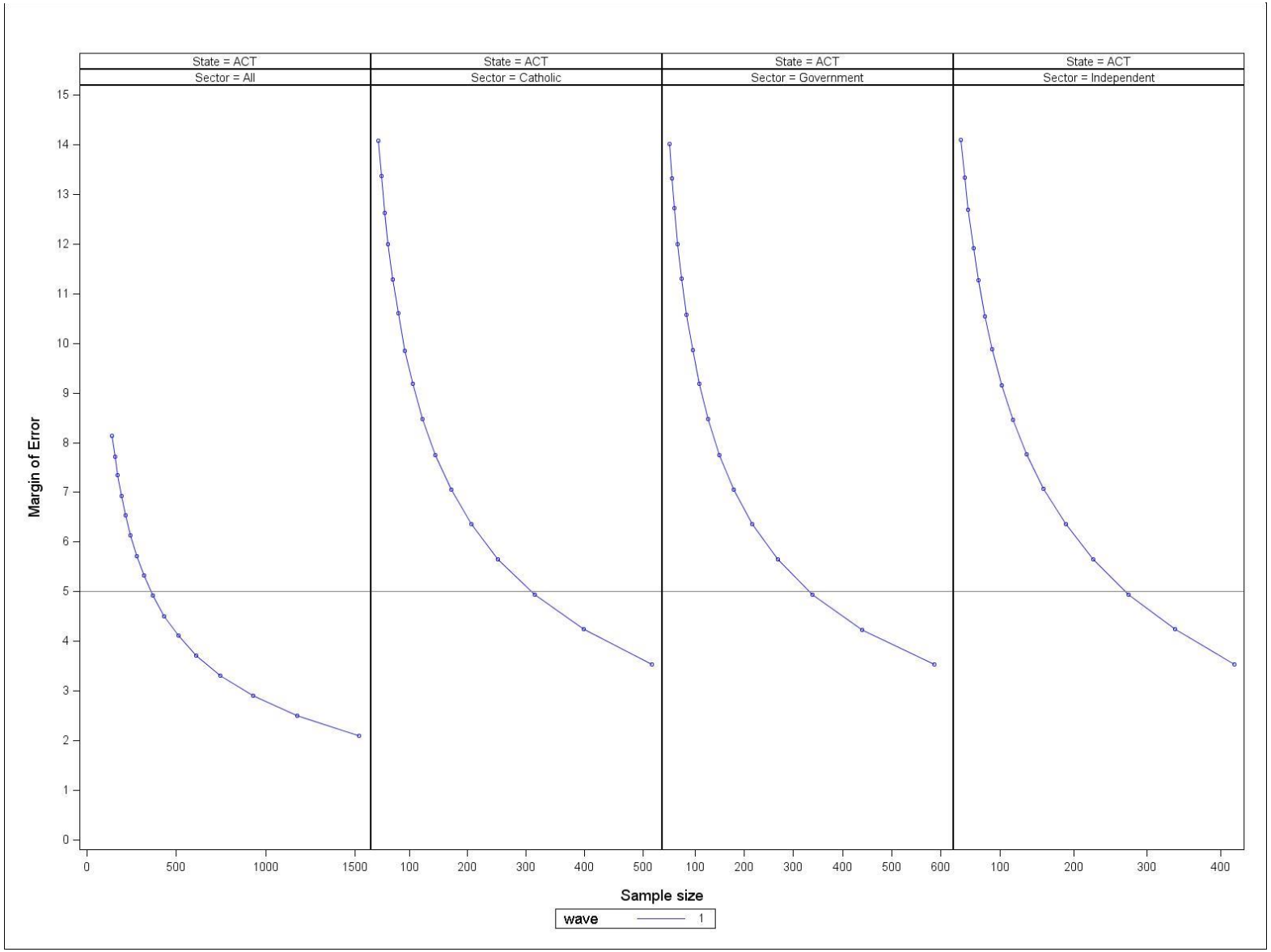


Figure 3 Margin of error for different sample sizes – NSW (assuming p = 0.5 and design effect = 2.0)

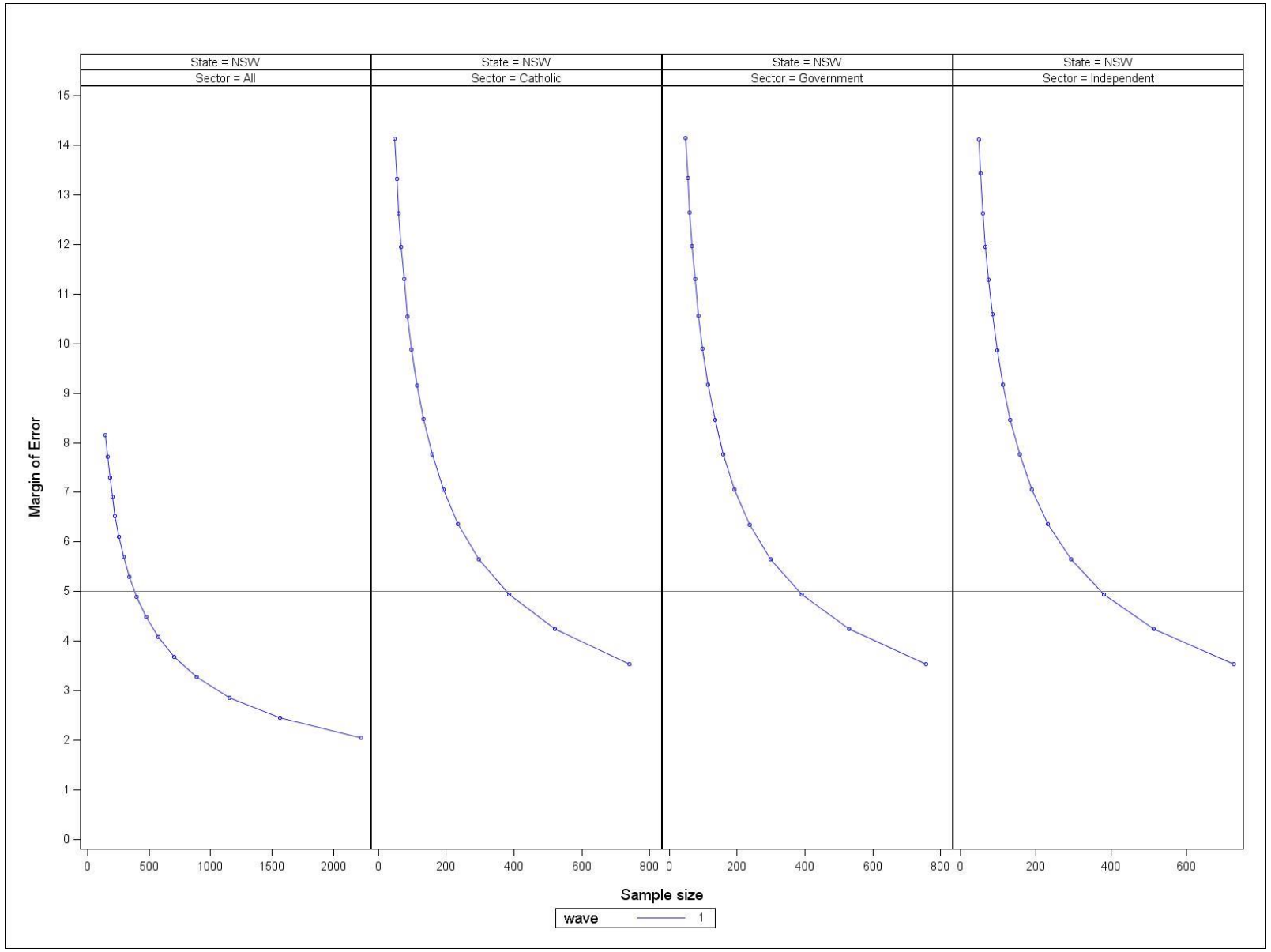


Figure 4 Margin of error for different sample sizes – NT (assuming p = 0.5 and design effect = 2.0)

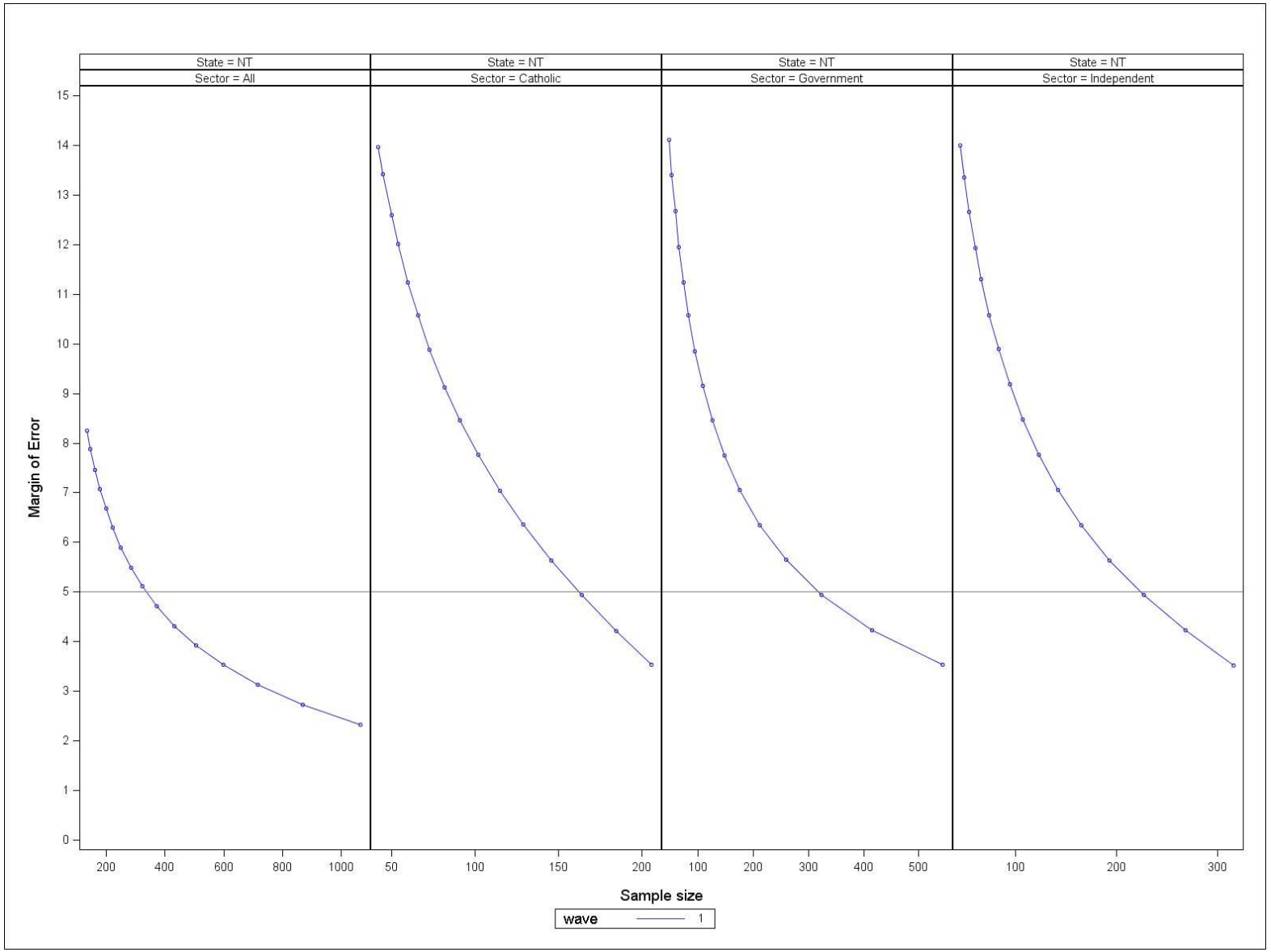


Figure 5 Margin of error for different sample sizes – Qld (assuming p = 0.5 and design effect = 2.0)

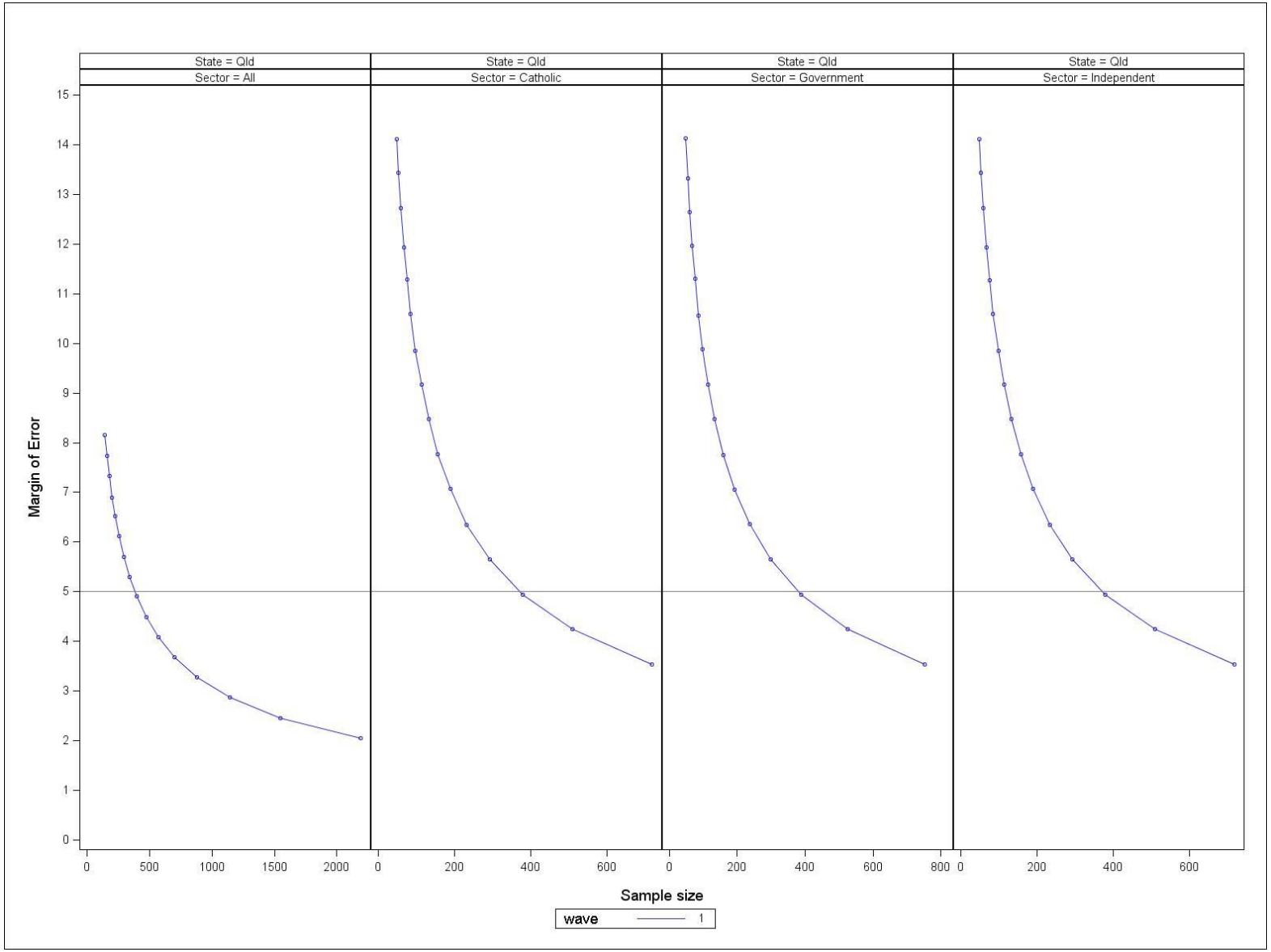


Figure 6 Margin of error for different sample sizes – SA (assuming p = 0.5 and design effect = 2.0)

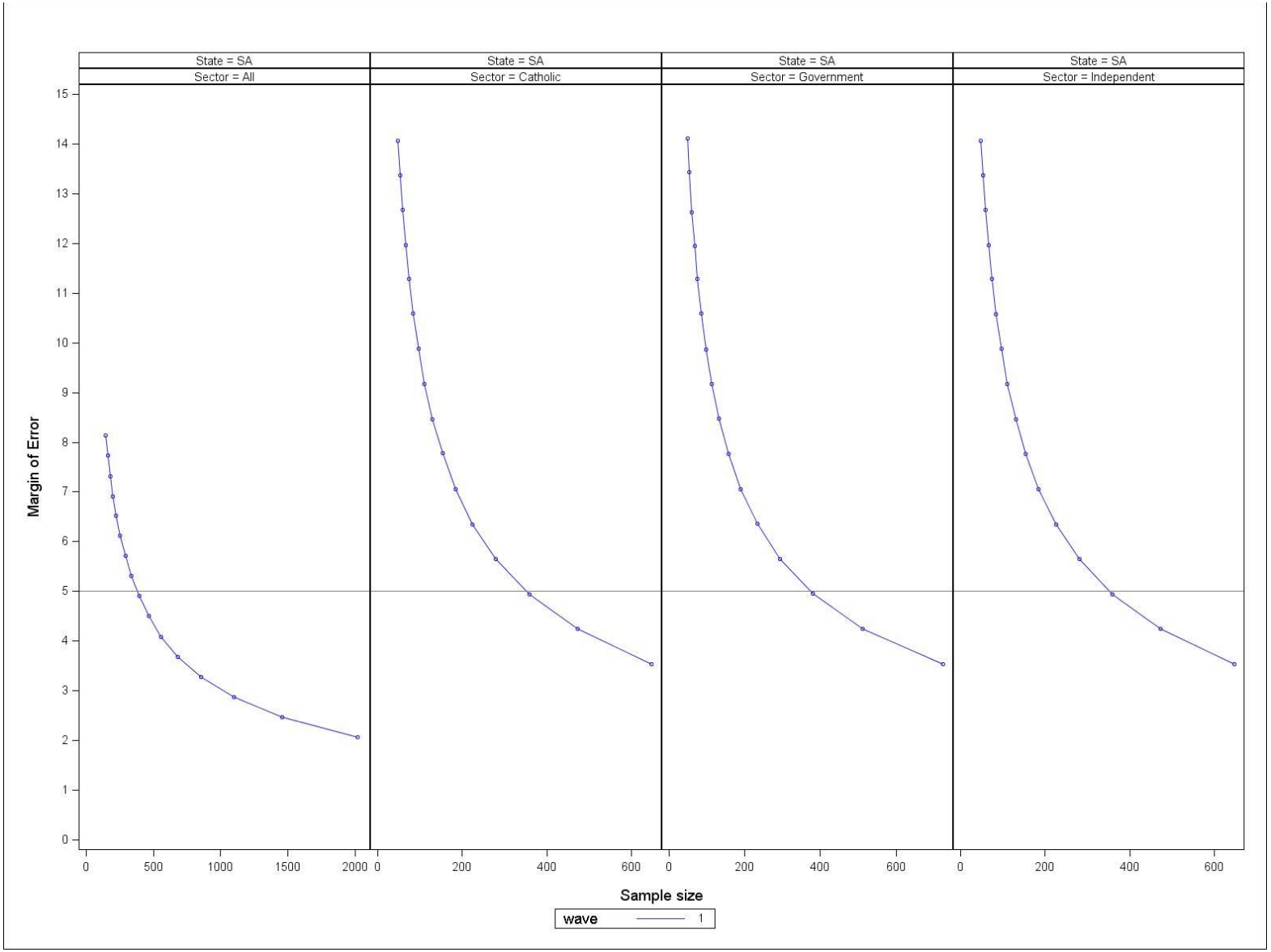


Figure 7 Margin of error for different sample sizes – Tas. (assuming p = 0.5 and design effect = 2.0)

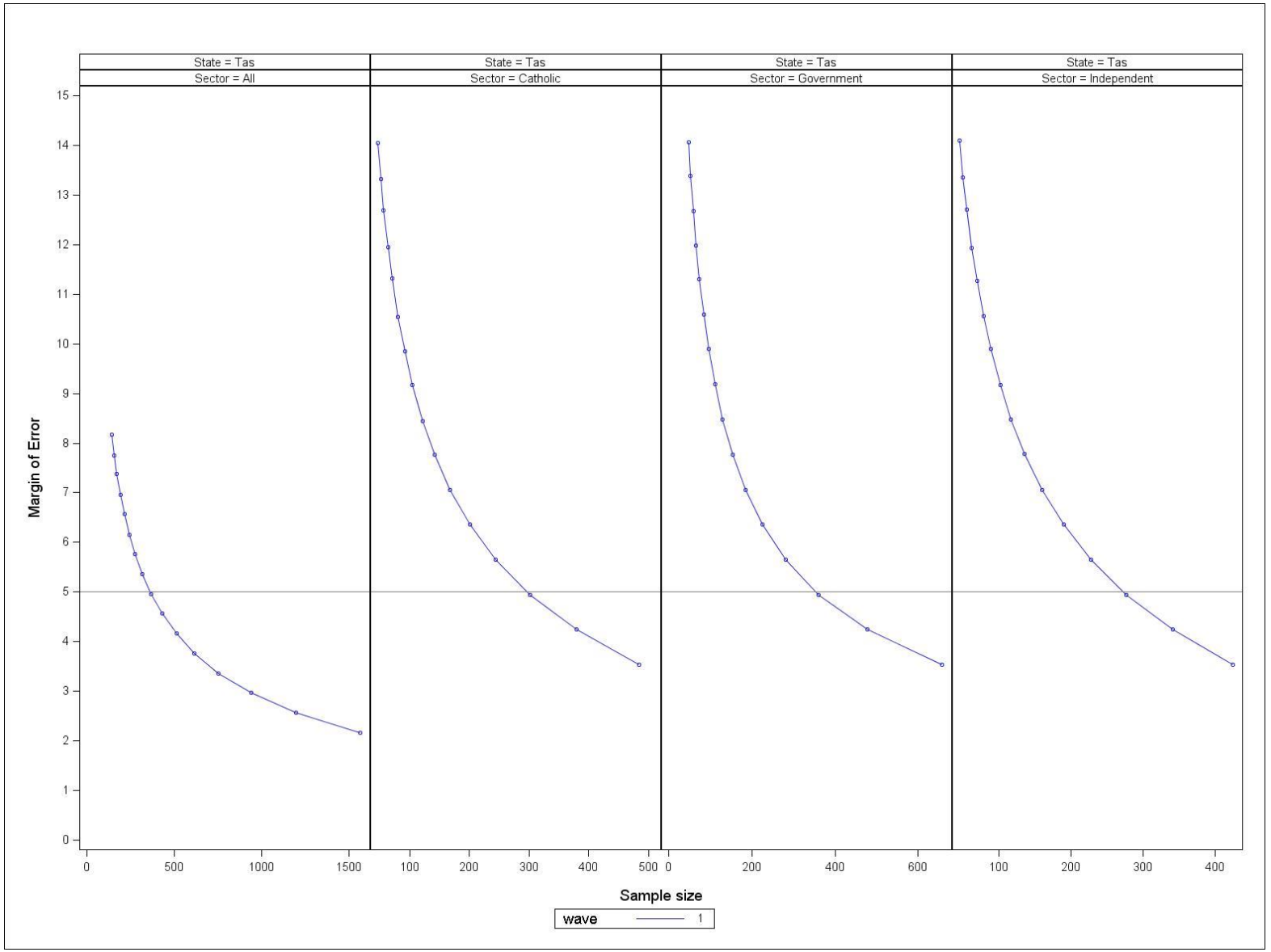


Figure 8 Margin of error for different sample sizes – Vic. (assuming p = 0.5 and design effect = 2.0)

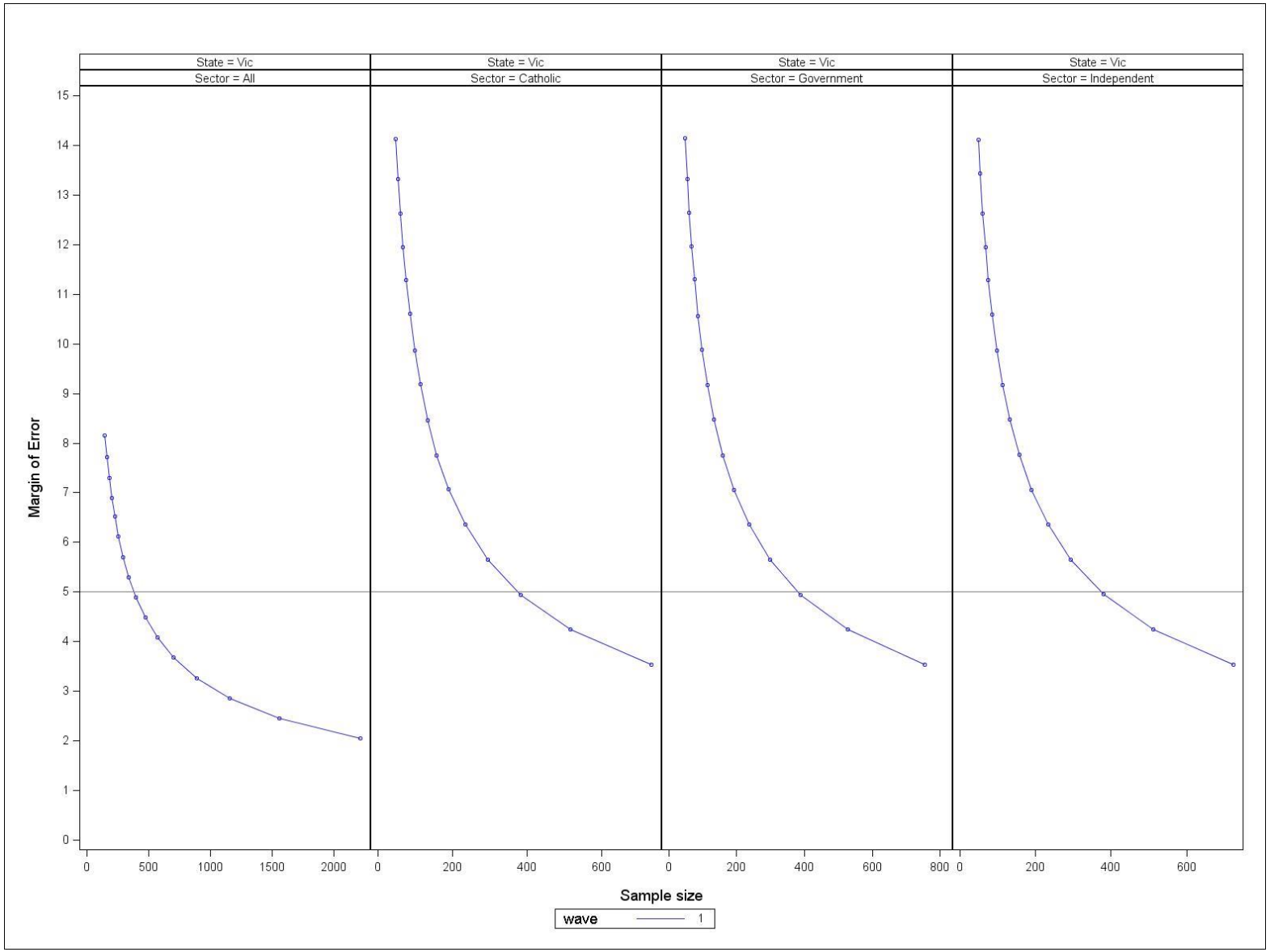
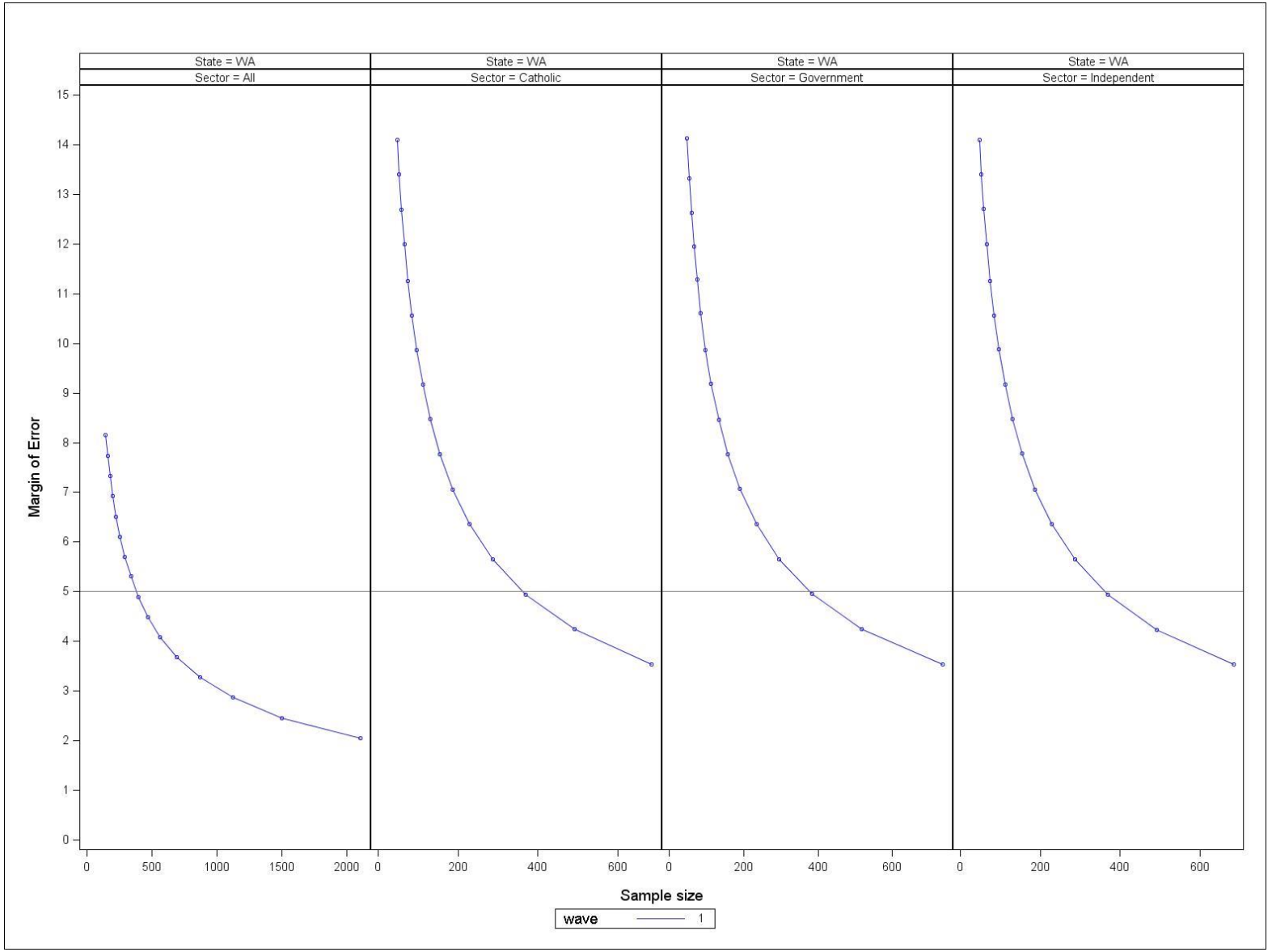


Figure 9 Margin of error for different sample sizes – WA (assuming p = 0.5 and design effect = 2.0)



Sub-populations

The value of LSAY would be increased if it possessed the ability to provide reliable data that measured the characteristics of important sub-populations.

The previous section showed that the current LSAY sample size provides robust estimates of population parameters for Australia as a whole, and for each of the designed state and sector combinations.

It is therefore also worth considering designing the LSAY sample to oversample (disproportionately sample) some of the key sub-populations. Further, these sample sizes need to also consider the impact of differential attrition over time.

Boosting sample sizes in any one sub-population may improve the accuracy (reduce bias) and precision of the analysis of this group. However, it can lead to inefficiencies when performing an overall analysis, or in other sub-populations (Collingwood et al. 2010; Citro & Kalton 1993; Kalton 2009). As Singh, Petroni and Allen (1994) note:

> Therefore, oversampling of a target group in a panel survey should be thought through very carefully before implementing. We are in no way suggesting that oversampling should be avoided in a panel survey. But, its usefulness should be evaluated in term of its long term effect on the goals of oversampling.

In determining whether to increase the size of sub-populations, the following factors should be considered:

- The sub-population needs to remain relevant for the life of the cohort.
- The population size of the sub-population needs to be able to be identified before drawing the sample.
- The total population of the sub-population needs to be large enough to draw an appropriately sized sample for robust statistical analysis.
- Are the sub-populations, time variant or invariant? In particular, consideration needs to be given to:
  - transition rates: the higher the transition rate, the lower the efficiency from oversampling
  - how long individuals remain in the sub-population
  - length of the survey (number of waves/years).

Sub-populations of interest may change over time, with individuals moving in and out of the sub-population of interest. Thus, sample sizes and precision estimates won't remain constant over all waves of the survey. Investment in ensuring good retention rates for these sub-populations needs to occur, as it is likely that these sub-populations will have higher attrition than the more general population.

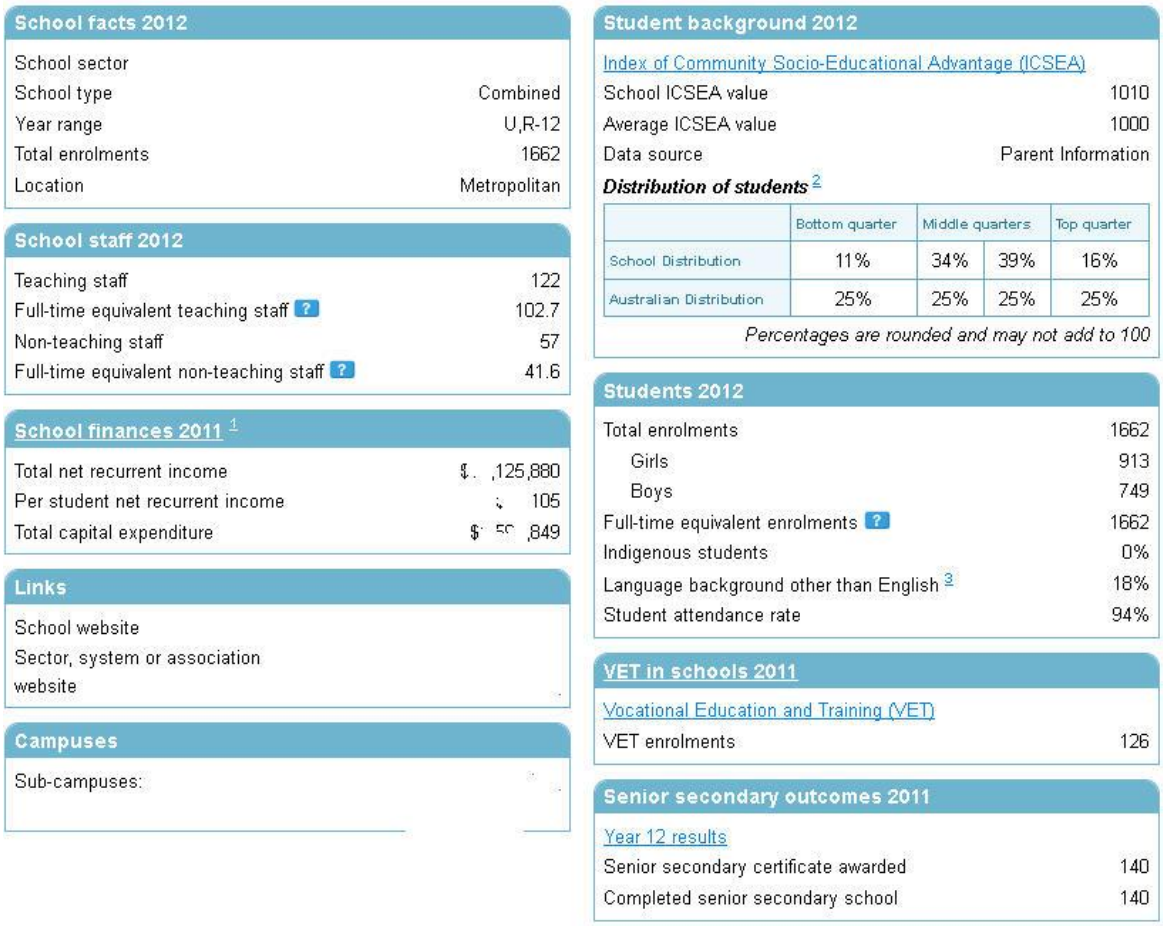
In relation to possible sub-populations for LSAY, it may be possible to further stratify the schools population into sub-groups of interest using data obtainable from the Australian Curriculum, Assessment and Reporting Authority (ACARA), where the MySchools website is compiled.

The information against which schools can be classified includes (see figure 10):

- socioeconomic status using the Index of Community Socio-Educational Advantage (ICSEA)
- overall school achievement in Year 9 NAPLAN (and earlier years for combined schools)
- gender mix
- percentage of Indigenous students
- percentage of students with a language background other than English (LBOTE)
- Year 12 completions
- size of school
- number of teachers
- student attendance rates
- school income information.

Using this information would enable the schools sample to be further stratified by school SES status, for example. The Australian Curriculum, Assessment and Reporting Authority data provide only high-level school information. Because student-level population information is not provided, its usefulness would only lie in sampling schools with the required characteristics. The students in the schools would still span the SES distribution and would still need to be chosen using simple random sampling. Obtaining individual population-level data on these types of characteristics for Australian 15-year-olds may be difficult; however, the data from ACARA may contain some of the required information.

Figure 10 School information available from the MySchools Website and the Australian Curriculum, Assessment and Reporting Authority



School facts 2012

School sector	
School type	Combined
Year range	U,R-12
Total enrolments	1662
Location	Metropolitan

School staff 2012

Teaching staff	122
Full-time equivalent teaching staff	102.7
Non-teaching staff	57
Full-time equivalent non-teaching staff	41.6

School finances 2011 [1]

Total net recurrent income	$,125,880
Per student net recurrent income	105
Total capital expenditure	$,849

Links

School website
Sector, system or association website

Campuses

Sub-campuses:

Student background 2012

Index of Community Socio-Educational Advantage (ICSEA)

School ICSEA value	1010
Average ICSEA value	1000
Data source	Parent Information

Distribution of students [2]

	Bottom quarter	Middle quarters		Top quarter
School Distribution	11%	34%	39%	16%
Australian Distribution	25%	25%	25%	25%

Percentages are rounded and may not add to 100

Students 2012

Total enrolments	1662
Girls	913
Boys	749
Full-time equivalent enrolments	1662
Indigenous students	0%
Language background other than English [3]	18%
Student attendance rate	94%

VET in schools 2011

Vocational Education and Training (VET)

VET enrolments	126

Senior secondary outcomes 2011

Year 12 results

Senior secondary certificate awarded	140
Completed senior secondary school	140

There are multiple sampling strategies (Kalton 2009, for example) that could be utilised to obtain oversamples for the sub-groups of interest; for example, screening or setting up multiple sampling frames. Should the option of removing the PISA link be considered, then these alternative methods could be further investigated.

As an aside, using an ACARA dataset has the potential to reduce the size of the school questionnaire, as some of the information asked of schools may be available from that administrative database. The ACARA datasets could be linked to the existing LSAY data as a possible enhancement.

In relation to these sub-populations, while the current LSAY—PISA arrangements are in place, it would be virtually impossible to modify the current sampling design without coordination between the LSAY and PISA Australian management teams and the OECD. However, under the existing arrangements, LSAY could adopt a screening approach to identifying and subsequently retaining individuals from key sub-populations in wave 2 (and subsequent waves). The responses provided by individuals in PISA (wave 1) could be used by the fieldwork contractor in wave 2 to ensure greater response rates for those in identified sub-populations. This could be achieved by employing some of the techniques outlined in the chapter on attrition (chapter 3). For time-varying characteristics, information would be obtained from the previous wave. However, an analysis should first be undertaken to ensure that the sub-populations targeted have a sufficient initial sample size.

Indigenous LSAY sample

In PISA, Indigenous youth are oversampled in an attempt to provide reliable estimates.

For this particular sub-population, the precision of the population estimates for 15-year-old Indigenous students using the current PISA sample sizes has been investigated. The 2012 schools collection (ABS 2012, table 40a) showed 12 695 Indigenous 15-year-olds attending school. This represents approximately 5% of all 15-year-olds attending school. Using the sample size calculations from the previous section, it has been assumed that Indigenous individuals comprise 5% of the samples (noting that this percentage was increased for oversampling, and then decreased due to higher attrition for this group).

A further assumption in the calculation of sample size presented in figure 11 and figure 12 is that the Indigenous population is spread across Australia in the same proportion as the overall population. The sampling scheme could be modified to ensure that the sampling proportions matched the distribution of the Indigenous population in each state or sector. That is, an alternative sampling frame could be established for surveying the Indigenous population.

Figure 11 and figure 12 (Australia and New South Wales only) show that the overall Australian estimates of the Indigenous populations in the early waves of LSAY are robust. At the national level, these estimates become less accurate as the sample size decreases due to attrition. For Australian school sectors (figure 11) and overall state (figure 12), the sample size of Indigenous 15-year-olds is not large enough to produce reliable estimates and any further demographic breakdowns for Indigenous youth will result in estimates with very large standard errors. It could be argued that the population of Indigenous 15-year-olds is too small to be properly represented in a national survey such as LSAY and might be better served through a separate longitudinal survey of Indigenous youth in a similar way that the Longitudinal Surveys of Australian Children (LSAC) and the Longitudinal Surveys of Indigenous Children (LSIC) are separate.

Figure 11 Margin of error for different sample sizes – Indigenous, Australia (assuming p = 0.5 and design effect = 2.0)

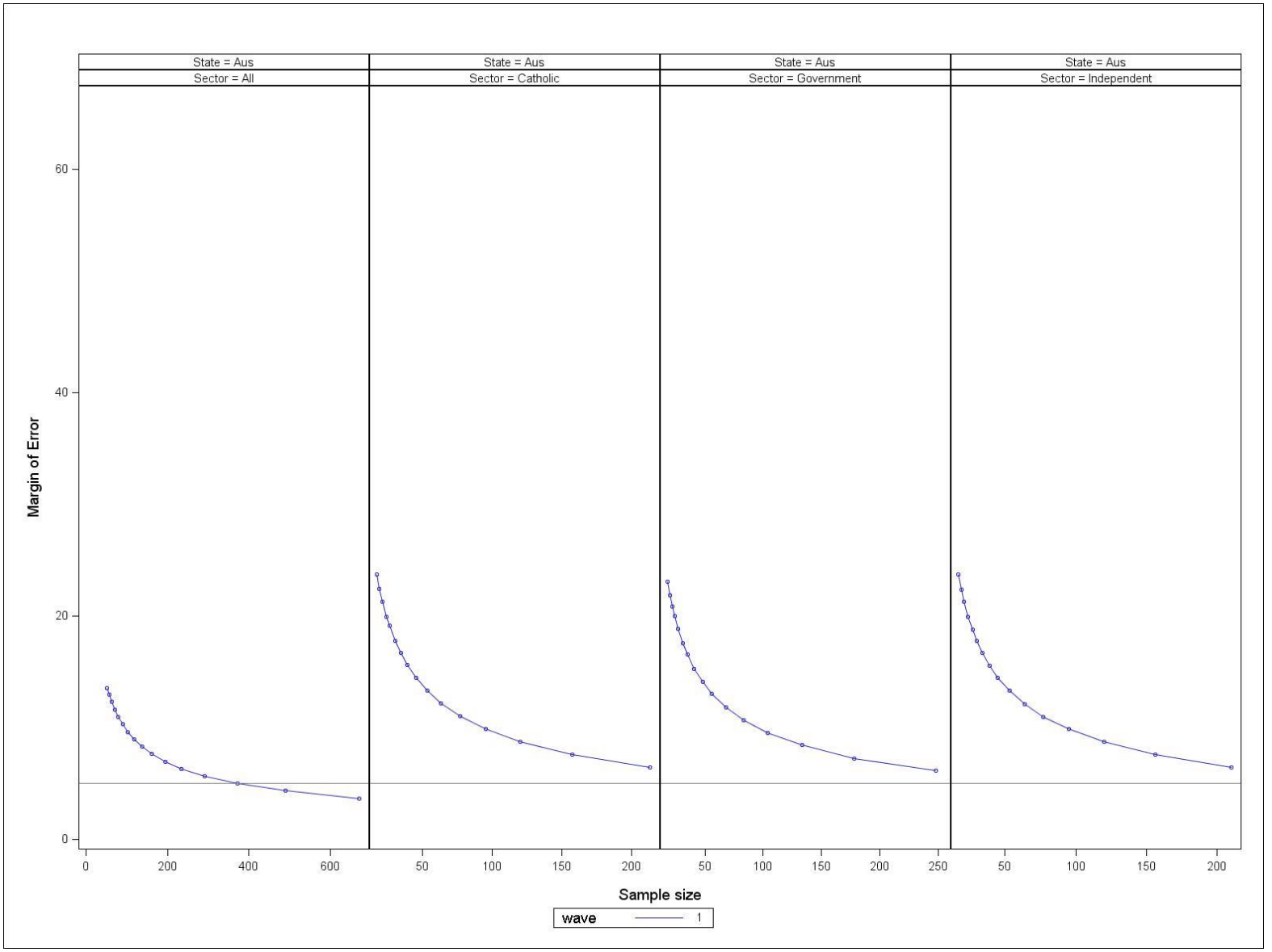
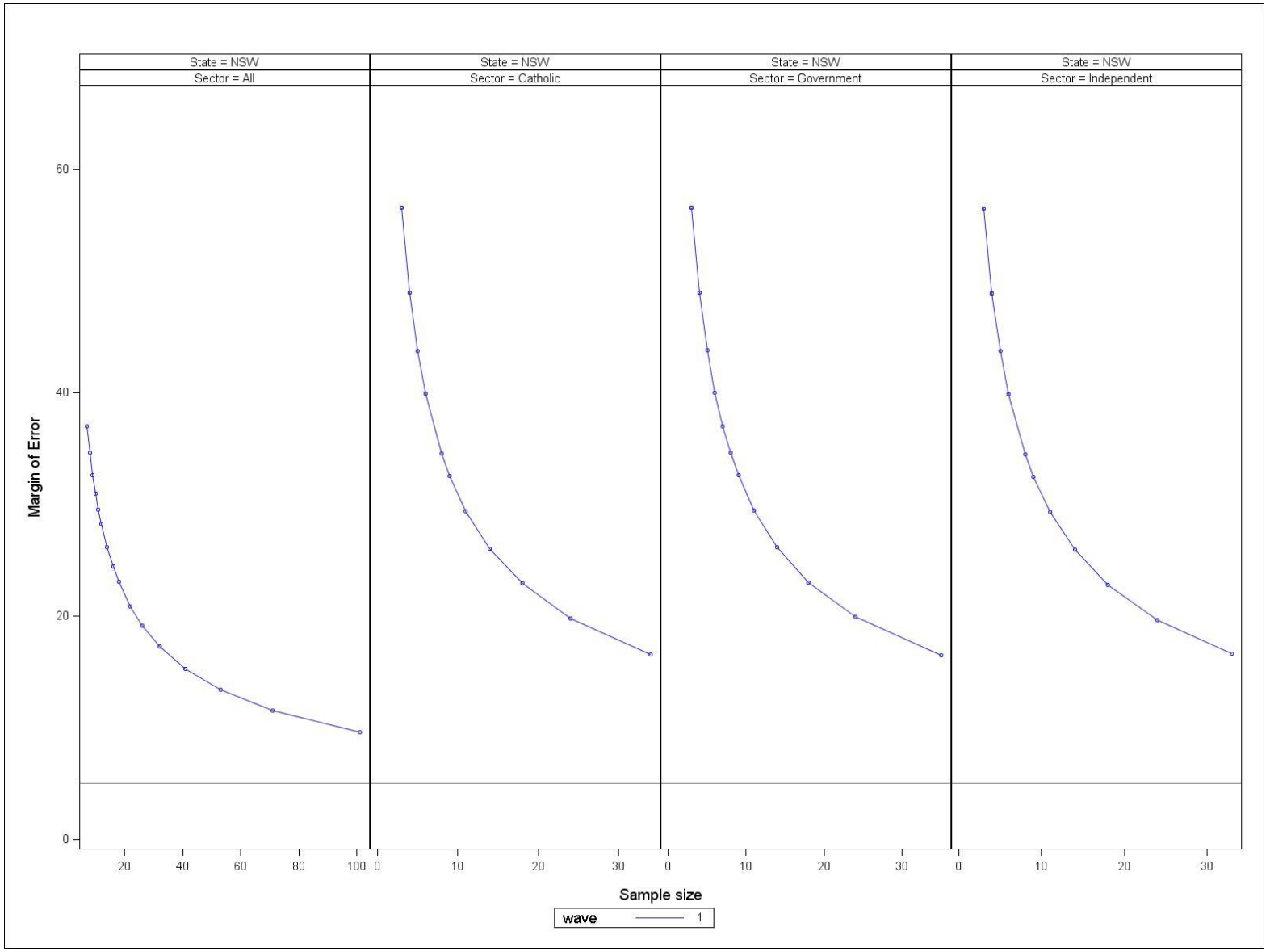


Figure 12 Margin of error for different sample sizes – Indigenous, NSW (assuming p = 0.5 and design effect = 2.0)



Schools

Under the current sampling strategy approximately 50 students from each school are sampled (OECD 2009). This results in around 300 schools being sampled (of a total of 2713), or around 11% of schools.

Table 1 presents the total number of secondary schools by jurisdiction and school sector. As can be seen, there is a small number of schools in some of the strata. In these cases, it would be worth considering sampling all Catholic and independent schools in the Northern Territory and in the Australian Capital Territory (and possibly all Catholic schools in Tasmania).

Further, if consideration is given to undertaking further stratification at the school level, this may necessitate choosing more schools to ensure adequate school sample sizes for the estimation of these sub-populations.

Table 1 Population of secondary schools by state and sector, 2012

	Sector			
State	**Government**	**Catholic**	**Independent**	**Total number of schools**
NSW	436	161	221	818
Vic	323	100	154	577
Qld	272	89	138	499
SA	144	33	56	233
WA	186	46	95	327
Tas	64	13	23	100
NT	88	9	17	114
ACT	27	*7*	11	45
Total	**1540**	**458**	**715**	**2713**

Table 2 Estimated number of sampled schools to ensure a sample size of 15 000 individuals, 2012

	Sector			
State	**Government**	**Catholic**	**Independent**	**Total number of schools**
NSW	15	15	15	45
Vic	15	15	14	44
Qld	15	14	14	44
SA	14	13	13	40
WA	15	14	14	42
Tas	13	10	8	31
NT	11	4	6	21
ACT	12	*10*	8	30
Total	**110**	**95**	**93**	**298**

Table 2 presents the estimated number of sampled schools for each stratum if 50 students are selected from each school. From this table it can be seen that more schools are needed than are available for Australian Capital Territory Catholic schools, and so the sample for ACT Catholic schools would need to be drawn differently.

In a cluster sample with relatively high intra-class correlations (that is, students from the same school are more alike than students from different schools), it is more useful to sample fewer students from more schools within each stratum (Cochran 1977; North Carolina Center for Public Health Preparedness; Aliaga & Ren 2006).

If the number of students per school were reduced to, say, 30 students (table 3), then the number of schools would be increased to close to 500 schools, an increase of around 200. This in turn would require consideration of the consequent increase in the number of school interviews and administrative arrangements. However, these costs should be evaluated against the possibility of gaining better sample sizes for the sub-populations under consideration.

Table 3 Number of schools sampled if 30 individuals from each school is selected (for sample of 15 000)

	Sector			
State	**Government**	**Catholic**	**Independent**	**Total number of schools**
NSW	25	25	24	74
Vic.	25	24	24	74
Qld	25	24	24	73
SA	24	22	22	67
WA	25	23	23	70
Tas.	22	16	14	52
NT	18	*7*	11	36
ACT	20	*17*	*14*	51
Total	**184**	**158**	**155**	**497**

From table 3, we again see that the school sample size for Catholic and now also independent schools in the Australian Capital Territory is greater than the number of schools available. The Northern Territory Catholic school stratum is close to the total number of schools. In this scenario, the sampling strategy for small strata needs to be investigated further. This scenario, however, would enable further stratification at the school level, as the greater number of schools would reduce the number of small stratum.

Cost considerations

The current sampling design produces accurate estimates at the national, jurisdiction and school sector levels, and consequently there may not need to be any changes to the design. There is scope to reduce the size for the larger jurisdictions and increase the size for the smaller jurisdictions. This would incur some costs to redesign the sampling strategy but would not result in additional data collection costs. There is also scope to reduce the number of LSAY participants in wave 1 (PISA) but any savings from reducing the sample size would need to be redirected towards improving attrition to ensure that the sample continues to provide robust estimates for the strata of interest. If the size of the sample is reduced, it will be particularly important to reduce the level of attrition for the Indigenous population. (Readers are referred to chapter 3 on attrition for a discussion on the strategies and costs to reduce attrition.)

Summary

The current LSAY sampling strategy and sample sizes are adequate for making robust population-based estimates for the strata of interest, even after attrition has been taken into account. The use of cluster-

based sampling is the most appropriate methodology and allows the use of multi-level models to investigate the impacts of school-level characteristics.

It has been shown that the sample of Indigenous students is too small to make robust predictions for anything other than at the national level, and this rapidly decreases as the sample size gets smaller, due to attrition.

While the LSAY program remains linked to PISA, there is very little scope to modify either the sampling sizes or the sampling strategy. Effort would be better spent in minimising attrition, particularly for the sub-populations of interest to policy-makers (including Indigenous). This might be achieved by considering some of the methods outlined in the chapter on attrition (chapter 3).

However, the severing of the link between LSAY and PISA would enable investigation of ways to make the existing sample more efficient; for example, by reducing the sample in some of the larger stratum and increasing the size of the smaller stratum. Further enhancement to the LSAY sample could occur by further stratifying the schools population into areas such as socioeconomic status etc. in order to select representative populations from these schools. This enhancement would significantly benefit policy-makers in terms of their being able to research issues relating to these sub-populations with confidence.

An alternative sampling strategy, whereby a smaller number of students was sampled from a greater number of schools, would assist in the stratification by these other school variables and also help to enhance the overall population estimates at all levels. However, this would entail increased administrative costs and administration of extra school-level questionnaires.

The LSAY sample size provides robust estimates of population parameters at the national and strata levels. Under the current arrangements, the PISA sample is controlled by the OECD and the Australian PISA team, both of which have strict design protocols designed to obtain robust national estimates for international reporting. However, there may be some scope to work with the Australian PISA team to make some changes to the current sampling scheme. Further, should this arrangement remain in place, resources should be spent in reducing the level of attrition from PISA to LSAY, and for the various sub-populations of interest to policy-makers.

3. Attrition in LSAY

The value of longitudinal youth surveys, particularly LSAY, is that they provide information and data on youth transitions and pathways that are not available using cross-sectional or time-series samples. They are designed to measure the relationships, outcomes and, importantly, the changes over time of the original survey population (in this case, the population of school-based 15-year-olds) after each wave.

However, like all longitudinal surveys, LSAY suffers from the loss of sample members. Due to its longitudinal nature, these losses are cumulative over time, which results in falling sample sizes. This, in turn, affects the precision of the research undertaken.

One of the limitations of LSAY (and most other longitudinal studies) is using them to estimate population figures (such as Year 12 completion rates) for years (waves) after the initial wave. Longitudinal surveys provide high-quality data for analysing trends and changes within the cohort of people surveyed over time. Generating population summaries beyond the first wave of the survey requires benchmarking against other available data sources and careful management of the survey to ensure that it continues to match the populations of interest (for example, the population of 17-year-olds). For LSAY, new entrants (migration) and people who leave the population (death, migration) would need to be considered for each year of the survey. In order to accurately capture the year-on-year populations, re-sampling and rebuilding of the new populations need to occur. While this is feasible, and there are many (such as household) longitudinal surveys that undertake this, it is an expensive process and needs reliable population (administrative or other) level data. Thus the use of longitudinal samples in providing age-based population estimates is limited. Given that the aim of LSAY is to represent the original population of interest, it important that the impact of attrition on the original population is minimised. By minimising sample loss in subsequent waves of the cohort, the collected sample should not deviate too far from the original population. This is particularly important if the loss of sample members is non-random. Non-random attrition is when different groups of people drop out of the survey at differing rates (for example, low-SES individuals may be more likely to drop out more than high-SES individuals).

The two main problems that attrition causes in the analysis of longitudinal data are:

- bias, that is, the sample no longer represents the original sample (noting that cross-sectional surveys also suffer from non-response bias)
- small sample sizes for sub-groups of interest.

Bias

A biased sample is one that doesn't represent the population of interest. For LSAY, this may mean that the proportion of young people reported as participating in higher education is higher than might be reported if all respondents remained in the survey.

The impact of attrition on bias is model-specific, that is, particular to the individual question being investigated. A sample may be biased for one estimate, but may be unbiased for a different estimate. It depends on the association between the variables from which the estimate is constructed. Previous research (for example, Fitzgerald, Gottschalk & Moffitt 1998; Jones, Koolman & Rice 2006) has shown that bias may not exist even if attrition rates are high. For LSAY, this may be the case when considering results based on the entire sample, that is, for Australia as a whole. However, when looking at smaller sub-samples, McCulloch (2001), Antonovics et al. (2000), Burkam and Lee (1998), and Watson (2003) have all found evidence of bias due to attrition.

Detecting bias in LSAY is not straightforward. It requires data that are not readily available from any administrative data collections (hence the usefulness of LSAY). The population under consideration is different from the cross-sectional information collected by agencies such as the Australian Bureau of Statistics. Ryan (2011), using LSAY data, showed that attrition appeared to increase slightly the estimates of Year 12 completions; however, there was no impact on the relationships between factors influencing Year 12 completion observed in his study.

Sample sizes

The second problem arising from attrition is that of a reduced sample size. Even if the survey drop-out is random (not linked to the characteristics of individuals dropping out), for certain sub-groups over a number of waves (years) the sample size may become too small to sustain robust statistical analysis. Often, it is these small sub-groups that lead to the more interesting results or comparisons (for example, regional, low-SES individuals). It is crucial for the important sub-populations to have adequate sample sizes in wave 1.

Attrition rates

From figures 13 and 14 which present the cumulative retention from PISA (or wave 1) and the year on year retention rates, we see that the LSAY year-on-year retention (figure 14) rate is comparable, (in the order of around 85%) once the respondents have been contacted or responded to the wave 2 survey. Comparatively, this is very good, particularly given the lack of paid incentive to participants.

Figure 13 shows the very high attrition rates for the Y06 and Y09 cohorts from PISA (wave 1) to wave 2. This is substantial (in the order of 20 and 40%) and greatly impacts on the overall sample sizes. It appears as though the attrition arising between PISA and LSAY is the result of both administrative arrangements and the behaviour of the young people. However, given that the response rates stabilise once the LSAY fieldwork contractors have made contact, it is likely that the administrative arrangements in place are contributing to this high attrition rate. In particular, part of this attrition is the failure of young people to provide correct or accurate contact details, and another part is due to the illegibility of the handwriting of the young peoples contact details.

The wave-on-wave attrition becomes more manageable and within the control of the LSAY management team when some the techniques listed later in this chapter are utilised.

Modifying the administrative arrangements, such as reducing the time lag between PISA and LSAY or matching administrative data to PISA respondents, may help to reduce the attrition. For example, the Y03 cohort maintained a high recovery rate from the PISA survey; this may have been due to the different methodology employed in wave 1 of LSAY, in which PISA participants were contacted again in the same year as they undertook PISA. In the 2003 cohort, young people were contacted via telephone to complete the LSAY component of the questionnaire.
If the LSAY—PISA link is maintained, then consideration needs to be given to greater coordination between PISA and LSAY to ensure that LSAY receives accurate student contact details following the PISA process. LSAY also requires the capacity to provide incentives to these students to encourage them to participate in subsequent (LSAY) surveys. Alternatively, running LSAY independently of PISA (such as using NAPLAN and developing its own wave 1[8] questionnaire) may assist in reducing this substantial early attrition.

[8] Wave 1 means in lieu of undertaking the PISA testing and questionnaire. Wave 1 of LSAY refers to PISA.

Figure 13 Cumulative retention rates for LSAY Y03, Y06, Y09, LSYPE, NLSY and the SWISS TREE Survey

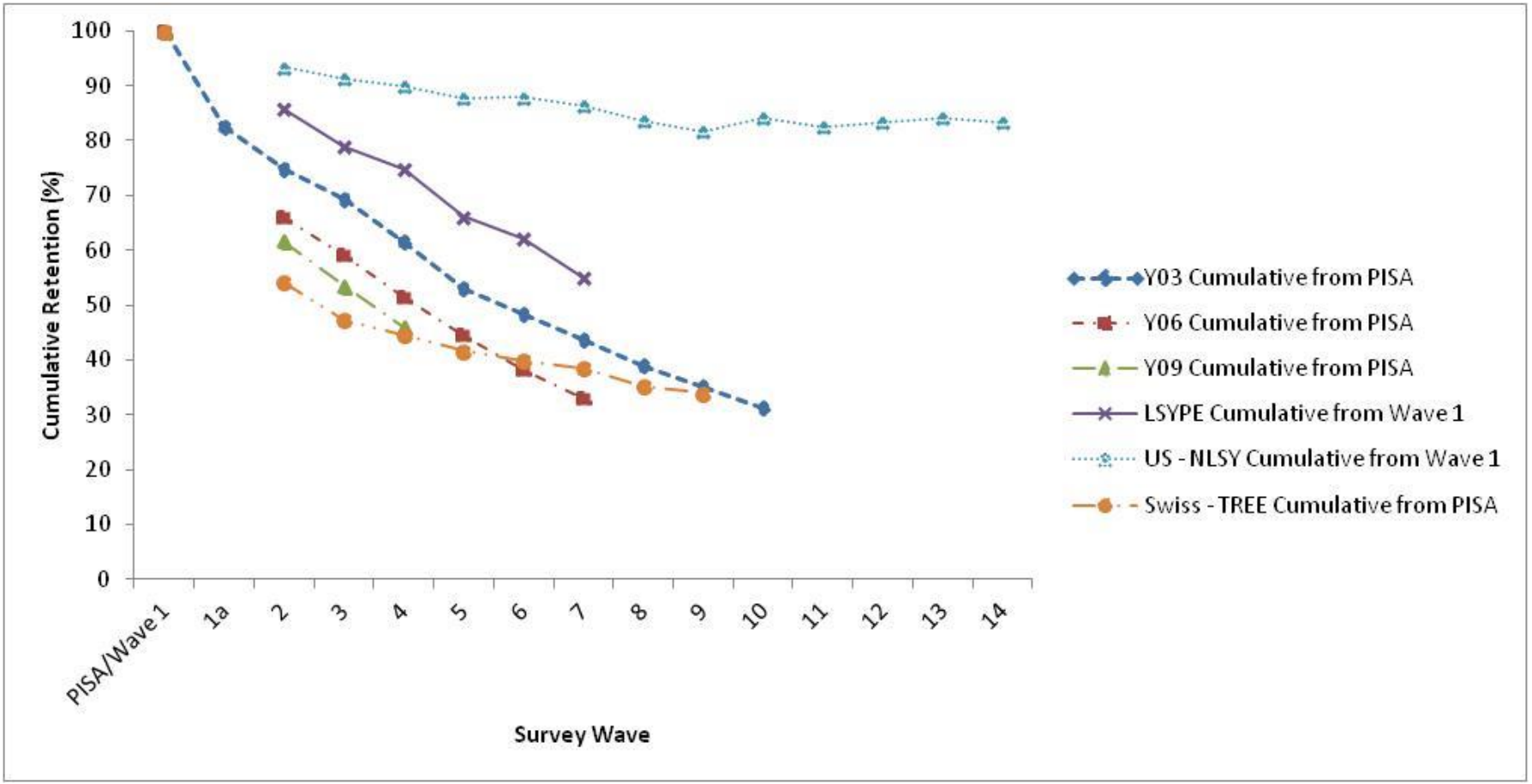
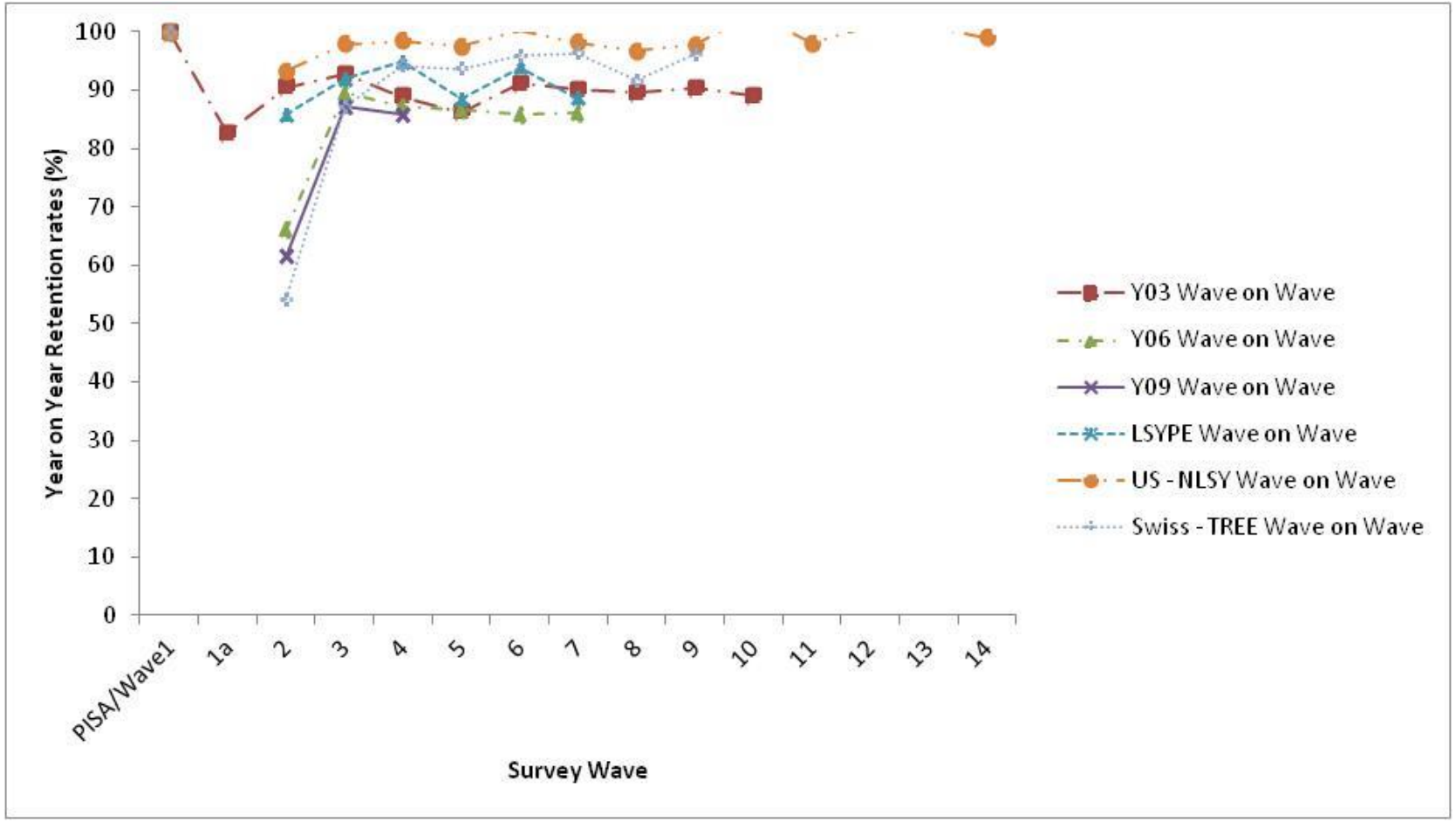


Figure 14 Year on year retention rates for LSAY Y03, Y06, Y09, LSYPE, NLSY and the SWIS TREE Survey



Year-on-year attrition accumulates substantially over the full timeframe of the survey. For example, by wave 10 for Y03, fewer than 50% of respondents remain in the survey. From figure 13 it can be seen that the overall retention of respondents in LSAY has been getting worse over time. This is driven by the increasing retention rates of respondents from PISA to LSAY, and slightly lower year on year retention rates for the Y06 and Y09 cohorts when compared to the Y03 cohort. Due to the nature of longitudinal surveys, the combined effects of these factors leads to a much lower sample size in the Y09 cohort.

Figure 13 also includes retention information from similar international surveys: the Longitudinal Survey of Young People in England (LSYPE); the Swiss Transitions from Education to Employment (TREE) and the US National Longitudinal Survey of Youth (NLSY). From this, it can be seen that, apart from the US NLSY, LSAY has year-on-year retention rates similar to LSYPE and slightly lower than TREE. The Swiss TREE survey also uses PISA as its first wave. TREE also suffers from significant attrition in the period between PISA and the first wave. However, TREE has higher year-on-year retention rates (greater than 90% after wave 2) and so has a better cumulative retention rate over the waves. LSAY Y09 and Y06 already have smaller retention rates than TREE, even though they have been in the field for a shorter period of time. The LSAY Y03 cohort is following a similar pattern to LSYPE.

It is worth noting that for LSAY, even with this level of attrition, the sample still represents some 4000 individuals. LSAY begins (in relative terms) as a very large sample of young people. Some similar longitudinal surveys begin with 5000 individuals or fewer as their original sample size and the issue of retention is very important. The US NLSY begins with a third fewer (10 000) individuals than LSAY. The NLSY is very successful in retaining sample members, the result of substantial follow-up, re-issuing members from the first wave, and very good incentives.[9]

Other surveys and in particular longitudinal surveys on youth are therefore suffering from attrition and at similar rates to LSAY. This problem is not unique to LSAY; however, the international surveys have slightly better year-on-year retention rates.

A more thorough investigation of the impact of attrition on key outcome variables is recommended, provided that reliable benchmarking datasets are available. These may include the jurisdictions' own schooling databases, the higher education and VET collections, ABS data and Centrelink and Medicare data. In particular, the real impact of attrition on the various sub-populations of interest should be investigated (for example, Indigenous Australians, noting that the PISA sample is representative of Indigenous Australians in wave 1).

In some regards, the problem of dealing with attrition should be left to individual researchers to determine the impact it has on their models and results. There are some robust statistical methodologies available to quantitative researchers that allow them to investigate and to attempt to deal with bias arising from attrition (and item non-response). The primary reason for this suggestion is that it would be practically impossible for a survey management team to adequately deal with attrition for every conceivable question put to the LSAY dataset. However, it is important that, where possible, attrition is minimised, particularly in the period between PISA and LSAY wave 1.

It is crucial that the LSAY data management team understand the potential impacts of attrition on the overall findings and estimates, particularly as they relate to sub-populations within the overall population. In order to ensure that the sample retains its validity over time (that is, match the original population),

[9] It would be nice to compare LSAY to the Canadian Youth in Transitions (YIT, <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4435&Item_Id=85022&lang=en#a3>) as this survey also used PISA for their first wave. The information isn't readily available from the Statistics Canada website.

retention rates need to be maximised. The next section will highlight some common methods for maximising retention rates in longitudinal surveys.

Dealing with attrition

Currently, LSAY relies on survey loyalty and an individual's sense of doing good to ensure reasonable response rates. There is evidence to suggest that this is a primary motivator to participation, particularly if young people know the study's purpose, who will use it, who has commissioned it and how their responses will be used to make decisions affecting them and their peers. However, with response rates declining, LSAY needs to find alternative methods to increase participation.

Incentives

Using incentives is one of the foremost ways by which participation in surveys can be encouraged. Incentives are used to maximise overall study response rates and can be targeted to specific groups within the overall population. They can take the form of money or vouchers, gifts, charitable donations or lottery draws.

In longitudinal studies, there is overwhelming evidence that the use of incentives increases response rates, especially for specific interest groups.

Any incentive needs to be of an appropriate level; that is, it needs to be relative to the cost of living and reflect the burden on respondents in terms of interview length. Experiments on the use of incentives have shown that prepaid (not tied to completing an interview) incentives produce the greatest impact on response rates. Incentives paid in the first wave were shown to increase response rates in later waves. In particular, unconditional incentives are useful in wave 1, as the aim is to obtain a large starting sample as possible (Berk et al 1987; Berlin et al 1992; Church 1993; James 1997; Mack et al 1998; Lynn 2001; Singer 2002).

Larger-than-normal incentives have shown to be effective in raising response rates in later waves. (In terms of LSAY, this might be the wave after the year that most of the respondents finish school.) Studies have shown that this incentive can then be reduced to normal levels in later waves without having a detrimental effect on response rates. It is possible also to modify, alternate or change the incentives without greatly impacting on participation on later waves.

A further use of incentives is for targeting hard-to-reach groups, or groups who are at a higher risk of dropping out of the survey. Larger incentives could be offered to the most difficult respondents to obtain cooperation and for those who may return to the survey after dropping out. However, there is an ethical issue, in that this is rewarding 'bad behaviour' more highly than good behaviour.

Incentives can also be used to ensure contact is maintained between waves; for example, by providing an incentive to keep contact details up to date, and to offer 'finders' fees' for family/friends who provide updated or new contact details.

Communication

Having a strategic marketing and communication plan with young people and their parents can help to reduce attrition. In a review of the Longitudinal Survey of Young People in England (Collingwood et al. 2010), a recommendation was made for including parents in the communication strategy. It was seen that parents can motivate and encourage their children to continue to participate in the survey.

Directly promoting the survey to young people and to those who may influence young people may reduce the loss of people from PISA to LSAY. This would include the development of a respondent-specific website, which should include the latest findings, media coverage, and a place for respondents to update their contact details and to claim incentives.

The development of an information pack for participants and their parents is another method of communication. This pack should include: an information booklet reinforcing the purpose, importance, and uniqueness of the survey; findings from previous cohorts; a membership card; a parental letter or pamphlet; and letters of consent (for example, parental permissions, data-sharing and access). The membership card should provide details on how to register online, details of the respondent website and how to update their details. The card may also include information on the study, contact details and a respondent ID number.

Additional communication strategies could be specifically built around young people in the 13 to 14-year-old age bracket to educate them about PISA and LSAY, as well as the 15 to 18-year-olds already participating to encourage them to stay in the survey. Their parents, schools and teachers should also be targeted.

Ongoing communication with respondents is seen as important, with items such as birthday cards or email and text updates (for example, when the survey is mentioned in the media) seen as a strategy designed to reduce attrition.

Communication strategies need to consider that young people want to know that this research matters and they are very keen to compare themselves with their peers. Young people are less concerned about issues such as privacy and data security, although it is important that their parents understand how these issues are managed.

Questionnaire

Survey length and complexity have been cited as reasons for people failing to respond. In general, longer interviews are related to respondents being less likely to cooperate in a future interview. However, very short interviews can reduce response rates as well.

It is important that young people find the questionnaire interesting and engaging.

Mode of survey

There are three primary modes for conducting interviews: personal one-on-one (computer-assisted personal interview — CAPI), telephone (computer-assisted telephone interview — CATI) and web-based approaches. In a longitudinal survey, there is scope to employ different methods for different waves and to use different approaches for different groups of people.

There is a general consensus that the first wave of a longitudinal survey should comprise face-to-face interviews. This approach enables concerns (either by parents, or the young people themselves) to be addressed early on, parental consent can be obtained and parents can be reassured with the survey. This approach would also enable the administration of a parental questionnaire if required.

As young people age and become busier, different modes could then be utilised. It is important that the mode of delivery is flexible and there may be times when a web or self-completion component may be more appropriate, particularly with sensitive topics. Self-completion is more reliable and is possible with the use of personal interviewing. By utilising a telephone interview, the respondent may never have to access the web-based questionnaire(s).

Web-based surveying is becoming more common; however, it is important to recognise that not all respondents have access to the web, and some may not like using it. Survey instruments sent via email can easily be ignored.

The advantage of a mixed-mode approach to surveying is that different modes can be targeted to different groups of people or individuals. For example, personal interviewing at a cafe or restaurant may encourage someone to respond who might otherwise not, or allowing respondents to undertake an interview when it suits them, either by telephone or the web, might be useful for busy people.

The use of mixed modes in a longitudinal survey may lead to potential bias: some questions asked via telephone may in fact be slightly different when asked via the web.

There is no evidence to suggest that interviewer stability influences response rates. Other strategies that could be employed include the use of life-course information when interviewing. For example, telephone operators could see the entire path of the respondent while on the phone. This approach helps the young person to remember their previous responses and also enables a reduction in item non-response. Other strategies include modelling the probability of response for each individual and providing that to the interviewer. This would be undertaken using information such as their own demographics, their previous call history and other ancillary information.

Re-sampling/reissuing

Other ways to minimise attrition without involving the respondents directly include rebuilding the sample or reissuing previously lost respondents.

The rebuilding of the sample involves re-sampling people from the original population frame or re-sampling sub-groups of the population (for example, new migrants). This approach is expensive and there are issues related to weighting and the representation of the population. Resampling from the original population frame in LSAY would require also administering the PISA testing. It may not be appropriate to administer the test to an individual who is older than the original cohort. Alternatively, if the PISA – LSAY link were removed, it might be possible to resample young people and obtain their NAPLAN scores.

The reissuing of respondents means that respondents who have previously been non-responders are placed back into the pool of people to contact in the current wave. Those who have previously given hard refusals are left out. With this strategy, effort will need to be expended in locating the individuals, particularly if non-response was related to non-contact. Reissues can be drawn from the original sample and in essence the sample is rebuilt from wave 1.

Both of these approaches require the use of a ‘catch-up’ set of questions, that is, a small subset of questions that will help to fill in the blanks while the respondent was not in the survey.

Technical and practical feasibility

Incentives

LSAY currently doesn’t offer any real incentives for participation. There is a random lottery draw for a $500 voucher, which can be spent at a wide range of retailers across Australia. There is a single prize for each jurisdiction for each cohort.

It would be a simple process to commence paying incentives to participants in LSAY. Targeted incentives could also be implemented using modelling as well as interrogating the data to determine the sub-groups

suffering higher-than-normal attrition rates. The only practical hurdle is that of cost, which in the case of LSAY may be substantial. Some of the costs of incentives could be offset by the reduced number of call attempts to survey respondents. The costs associated with incentives also need to be considered when reviewing the frequency of surveys, how often a new cohort commences, the mode of survey delivery, the inclusion of a parent and/or a teacher interview, the number of waves, and overall sample size. Further, the cost of using an incentive should also be considered when investigating the addition of separate modules, particularly if they relate to sensitive topics (such as health and wellbeing, addressed in chapter 5). The cost of using an appropriate incentive may in fact be offset by the use of a smaller initial sample.

If no provision is made for an appropriate incentive, future cohorts of LSAY will continue to suffer from low response rates (as is already being experienced with the Y09 cohort).

As noted in earlier chapters, in LSAY a substantial issue emerges in the level of attrition that occurs in the period between PISA to LSAY (wave 1 to wave 2). Although this attrition can be attributed to several factors, it is primarily due to the lack of accurate contact details supplied by the young people when they undertake the PISA survey.

A key point from the discussion earlier in this chapter was the effectiveness of incentives paid in the first year of participation. This incentive can influence the decision to continue with subsequent waves of the survey. Given that LSAY is linked to PISA, potential incentives are available only to those who participate in wave 2. The response rates of the second wave may be improved if an incentive was given to PISA participants if they provided reliable up-to-date contact details at the time of PISA. That is, once the LSAY fieldwork contractor had the contact details and made initial contact, an incentive (along with an LSAY information pack) could be provided to the young person. This approach would require the cooperation of the Australian PISA team, the OECD and all school jurisdictions, as it would require advertising during the PISA surveys.

Alternatively, consideration could be given to sending unconditional incentives (along with a survey information kit) to the young people and their parents soon after they have sat PISA. This would require the contact details of PISA participants to be passed to the LSAY team fairly soon after completing the PISA test, necessitating more effective information-sharing between the Australian PISA team and LSAY.

A final approach to obtaining better contact information might be through the schools themselves. Providing financial compensation to the participating schools may encourage them to thoroughly check the contact details of their students who sat the PISA testing (along with an administration fee to the Australian PISA team).

Communication

The level of communication with LSAY participants is already fairly high. There is a regular six-monthly sample maintenance program that comprises a ‘newsletter’ and a letter/postcard in the periods between interviews. The newsletter has evolved over time to be youth-focused, with the content highlighting issues relevant to young people. Recently the newsletter doubled as a YouTube video clip. However, the impact of the current communication strategy is unknown, and it would be worth considering undertaking market research on the impact of the sample maintenance products on LSAY participants.

The production of a respondent-only webpage would be feasible and it could focus on providing snapshots of the data, relevant media coverage of LSAY and other respondent-focused content. An important aspect of this website should be enabling the young person to compare their own situation with that of their

peers. The existing LSAY website is an effective tool, but it is primarily aimed at the researcher or data user and is not appropriate for respondents.

The development of an information pack and other supporting material is feasible, and would have maximum effect if delivered at the time of the PISA test, or at the very least prior to the wave 2 telephone interviewing. The contents of this information pack would need to be addressed to both the respondent and their parents, the aim being to ensure consistent and ongoing participation.

Additional general marketing of LSAY to young people would be more difficult and would require the survey managers to have general, mass media coverage to the young people at the right time (that is, a year before the first wave commences).

Further consideration should be given to maintaining ongoing communication with survey participants and could take the form of email and/or social media. The alert or communication could be triggered when LSAY featured in the media or another significant product using LSAY was released. Alternatively, a regular (monthly) email, tweet or other update could be sent to LSAY participants to maintain regular contact and to demonstrate to them how the data they have provided are being used.

Questionnaire

The questionnaire is one area where a significant impact is possible. Respondent fatigue and burden is the primary reason for young people dropping out of the survey.

The nature of the data that LSAY is attempting to capture means the content can be unappealing and seemingly irrelevant to the young people completing the questionnaire. By investing more time in questionnaire development, planning the questionnaire over the life of the cohort and utilising a youth-focused questionnaire reference group,[10] a questionnaire that is interesting and engaging for young people could be developed.

In a more structured approach to questionnaire and cohort design, alternative delivery methods (such as web-based) can be planned and assessed and include sensitive topic areas. Similarly, the inclusion of new areas, such as health and wellbeing can be fully integrated into the questionnaire without length being compromised.

Although the current questionnaire hasn't been modified significantly over the course of the current LSAY program, a disadvantage arising from its modification would be that cohort comparisons may become more difficult. However, the benefits of a more engaging questionnaire and one relevant to the experience of respondents as well as its potential impact on attrition would counteract this one negative.

Mode of survey

The current mode of delivery, the choice of CATI or web-based surveying, is appropriate for the older members of the cohort, particularly since respondents have the option to use either. It would be useful to investigate whether differences in responses arise between telephone and web-based methods. A further investigation should be undertaken to determine whether there are differences in the characteristics of respondents who use the different survey modes.

[10] A youth reference group comprised of young people spanning the age of the LSAY cohort could also provide advice on sample maintenance and other products intended for the respondents. A group such as this is being successfully run in the United Kingdom. Further details can be provided upon request.

As a method to help reduce attrition, LSAY could ask respondents for their preferred mode (web or phone) in one wave for use in the following wave. An alternative mode could then be employed if they do not respond within a given timeframe, using their preferred mode. It is possible that mode choice could result in soft refusals, particularly if the web option is selected.

The approach taken in the Y03 cohort, where respondents were contacted in the same year of PISA (2003) for a follow-up CATI, resulted in a lower attrition rate from PISA to LSAY. If the PISA link is to be maintained, then it would be worthwhile considering re-adopting this method of interviewing in the first wave, along with the communication and incentive strategies previously discussed.

If the decision to undertake an alternative first wave for LSAY is adopted, it would be worth considering a change in mode to face-to-face, incorporating both a parental and respondent questionnaire. However, at this stage the implementation of this is dependent on the outcomes of the review.

It would be feasible to calculate the probability of response for each individual in LSAY. This information could be made available to the fieldwork contractor for inclusion in the information on each respondent.

Influence on analysis and reporting

A general increase in response rates would decrease the chance of bias in reporting. However, the real benefits would arise if the increase in response rates results in a greater number of the sub-groups remaining in the survey (for example, Indigenous).

Cost considerations

It is difficult to estimate the costs for adoption of all of these possible enhancements for reducing attrition. Estimated costs are provided for some of the enhancements, based on consultations with survey experts and a review of the literature. However, the methods adopted and associated costs need to be considered in conjunction with any other changes recommended by the review panel. For example, a higher response rate could result in a smaller beginning sample, thereby offsetting some of the costs of providing an incentive.

Incentives

The cost of implementing incentives will be high but will vary depending on when incentives are offered, the level of the incentive and the number of eligible participants (all participants or a targeted subset). Given that there is currently a high level of attrition from PISA to LSAY (wave 1 to wave 2) one approach is to offer incentives to all participants if they provide up-to-date contact details at the time of PISA. Under this approach, the cost of providing an incentive could range from $300 000 (based on 10 000 participants and an incentive of $30[11]) to $420 000 (based on 14 000 participants and an incentive of $30). There will also be costs associated with the management of this payment. As mentioned, some of these costs will be offset by lower call attempts to survey participants.

[11] The level of incentive needs to be of an appropriate value. For the purpose of estimating costs, a level of $30 was applied. This is based on the level of incentive offered in the Household Income and Labour Dynamics in Australia (HILDA) survey. Watson and Wooden (2012) indicate that HILDA participants in wave 13 will receive an incentive $35.

Survey mode change

The approach taken in the Y03 cohort, where PISA participants were contacted in the same year of PISA for a follow-up CATI, resulted in a lower attrition rate from PISA to LSAY. Readopting this method of interviewing in wave 1 will involve additional fieldwork costs but it is estimated that the length of the interview would be around half the length of interviews in waves 2 and onwards. PISA costs will also decline by approximately $40 000 through the removal of the LSAY questionnaire at the time of sitting PISA.

Possible survey modes for a wave 1 interview include personal interviews (CAPI mode) or telephone interviews (CATI mode). Estimating the differences in costs between CAPI and CATI is difficult but there is a general consensus that face-to-face interviews have significantly higher costs than telephone interviews but result in higher response rates (see for example Richardson, Ampt & Meyburg 1995).

Doyle (2005), Groves and Kahn (as cited in Weisberg 2005, p.295) and Ballivian and Azevedo (as cited in OECD 2013) suggest that the cost of personal interviews is at least twice the cost of telephone interviews. Dillman (as cited in De Vaus 2002) and Groves et al. (as cited in Weisberg 2005, p.295) suggest the total cost of personal interviews could be five times that of CATI mode. Consultations with survey experts confirm that the cost of personal interviews could be five times the cost of conducting telephone interviews when all additional costs are considered.

The cost difference between the modes arises because the cost of interviewing is greater for personal interviews compared with telephone interviews due to the higher rate of charge for face-to-face interviewers and the greater time required to interview. Doyle (2005) estimates that the length of personal interviews is three times that of telephone interviews. Costs associated with interviewer travel (including accommodation, travel and time lags between interviews) also make personal interviews considerable more costly (Richardson, Ampt & Meyburg 1995; Doyle 2005; Kaminska & Lynn 2013). Groves and Kahn (as cited in Weisberg 2005, p.295) estimate that the length of a half hour face-to-face interview translates to approximately 8.7 hours of interview time for a national survey when travel is included, compared to 3.3 hours for an equivalent half hour telephone interview.[12]

There are other significant costs associated with face-to-face interviews such as the organisation of interview times and the redesign of the sampling frame to allow a more manageable clustering of participants (Richardson, Ampt & Meyburg 1995; Doyle 2005; Kaminska & Lynn 2013).

Communication strategies

Improved communication with participants and their parents can occur through several mechanisms. As noted, the development of an information pack sent to participants and their parents at the time of the PISA test is feasible. The cost to project manage, develop and design a double-sided DL sized flyer is estimated to be around $20 000. Additional costs will be incurred for colour printing (approximately $3000 for a single flyer at a quantity of 16 000).[13]

The development and printing costs will need to be multiplied if several promotional products are included in the pack. Costs will also increase if a personalised product is produced, such as a membership card with each participant's name or LSAY identification number. The costs will vary depending on

[12] These figures may now be slightly outdated due to the lower costs of long distance calls today and the current use of computer assisted interviews (CAPI and CATI modes).

[13] This costing is based on expenses incurred for the production of a thank-you card for the 2013 LSAY sample maintenance program.

whether a web-based approach or a mail approach is implemented to send the information pack. The mail option will have additional costs related to printing, postage, envelopes, and mail management.

Summary

Attrition in longitudinal surveys is complex and difficult to manage. There is no single remedy that will increase survey response rates. Further, there is no guarantee that the suggested strategies will have any substantial impact on the response rates of LSAY, although failure to investigate the viability of these remedial strategies could lead to reduced sample sizes, which would diminish the impact of the LSAY data and its relevance and usefulness to policy-makers.

An area that needs immediate investigation is the higher-than-expected attrition rates in the period between PISA and the first LSAY wave. It is important that the reasons for this are investigated, with any possible remedies implemented as a matter of priority. The high level of attrition may in fact be due to the administrative arrangements rather than the behaviour of respondents. An investigation of these arrangements might therefore offer insights into the cause and therefore offer a remedy for raising the PISA/LSAY retention rates. This would require the cooperation of the LSAY management team, the Australian PISA team, and possibly the OECD PISA managers and relevant school authorities. Any strategies may require changes to the existing methodology of PISA.

It is important not to dismiss the importance of the role that parents can play as advocates and motivators of their children to continue to participate in LSAY.

Further investigation into the true impacts of attrition on the estimates obtained from LSAY and whether the overall relationships and estimates are impacted by attrition are necessary. The use of survey and attrition weights can go someway towards addressing the impacts of attrition. However, the value and impact of weights diminishes as the questions and models include variables that have not been used when undertaking the weighting, or when the mechanisms that influence attrition for a particular problem are not considered in the supplied weighting variables.

The question of whether sub-populations are impacted by attrition will depend upon whether the sub-populations were a representative sample in the original sample (PISA).

A clear statement of the aims of LSAY is required. Dealing with attrition could be an expensive proposition, particularly if incentives are included. It is important to define what the relevant stakeholders want from the data. The costs associated with improving the measurement of pathways, trends and outcomes for young people on a national level will vary according to the level and extent of detail of the information demanded from the survey.

4. Extending LSAY beyond age 25

The OECD (1996) describes a successful transition into adulthood as:

> The passage from adolescence to adulthood was historically signified by several, sometimes distinct, sometimes simultaneous, steps. These include leaving one's parents and setting up one's own home, forming a couple, more often than not marriage and a family, and settling into a more or less stable job, which is often an important catalyst for the other steps.

Even though the transition from youth to adulthood is becoming longer and more complex, the underlying indications of a successful transition remain the same. These indicators include:

- finishing full-time education
- settling into stable employment
- moving out of the parental (or primary care givers') home and setting up new living arrangements (ABS 2009)
- forming close personal relationships outside the immediate family.

The description given by the OECD (1996) implies the passage of time, rather than a definitive crossing of a given point. This time point differs for different individuals, but it can also change for cohorts as a whole due to general social and economic trends.

For example, the changing nature of the labour market, governmental policy and general economic conditions all influence whether young people acquire full-time stable employment, at what age they are able to leave school, and whether they can afford to leave the parental home.

In any longitudinal survey involving youth transitions it is important that the end point be at a time when a majority of young people have made the transition from youth to adult. Currently LSAY ends when the participants are around 25 years. This chapter provides some evidence that the length of transitions for young people is growing longer, and describes how LSAY could be enhanced so that the end point of the survey captures this transition.

Increasing length of transitions

An increasing amount of evidence shows that youth transitions are becoming more complex and consequently taking longer.

Education and employment

It is becoming more common for young people to undertake higher education when they are older (for example, by taking a gap year or through alternative entry) (Bynner et al. 2002). Young people also take multiple and complex pathways through their education and employment journey. Often, the two paths are interconnected, particularly when young people undertake traineeships as part of their part-time jobs (particularly in the areas of retail and fast food). Furthermore, the introduction of VET in Schools programs and school-based apprenticeships are further encouraging young people to stay at school longer.

Government policy that aims to increase participation in post-school education sees more people undertaking higher-level courses and therefore taking longer to complete their courses. Further,

policies such as increasing the age at which young people can leave school and 'earning or learning' requirements are protracting transitions, with young people now required to remain in education until age 17. This may have lengthened the time that it takes young people to finish their post-school courses (Smyth, Zappala & Considine 2002).

The prevailing economic conditions and cost-of-living pressures have seen the proliferation of young people working while studying. Thus, in terms of indicators of a successful transition, short-term employment indicators don't necessarily reflect genuine longer-term outcomes.

Further, permanent structural changes to the labour market, particularly the casualisation of the labour market and the reduction in the numbers of low-skill jobs, have given young people less stability and certainty of future employment. With the labour market moving away from the model of stable, permanent employment, young people entering the job market today are unlikely to find a job for life and are likely to have a high degree of job mobility (Smyth 2002).

In terms of higher education graduates, Purcell and Elias (2004) show that graduate pathways evolve slowly, and some graduates take five years or longer to settle into their careers. These pathways often involve further study, false starts or a rethink of their early career choices.

Figure 15 presents the percentages of the Y95, Y03 and Y06 in full-time study and full-time employment over the life of the cohorts (ages 14 to 25).

From this figure, it can be seen that, even over this relatively short time period, there has been an increase in the age (indicated by shift to the right) at which the young people switch from mainly being in full-time study to mainly being in full-time employment. For the Y95 cohort, this crossover point was around the age of 21; for the Y03 and Y06 cohorts, this has moved to the right to be at around 21.5 and 22 years respectively

A key definition of transition is the point at which the majority of young people are not in full-time education (Ainley, Malley & Lamb 1997; Smyth, Zappala & Considine 2002). In terms of this key indicator, figure 15 shows that for the Y95 and Y03 cohorts, by the age of 25, fewer than 20% of the cohort remains in full-time education and more than 60% are in full-time employment. However, this indicator is rather coarse, in that it doesn't recognise the type of employment in which young people are participating (for example, occupational categories, employment that matches qualifications and skills, nor the softer measures such as satisfaction with their employment), which means that the complexity of the transition process may be masked.

Figure 15 Percentage of cohorts undertaking full-time study, full-time employment

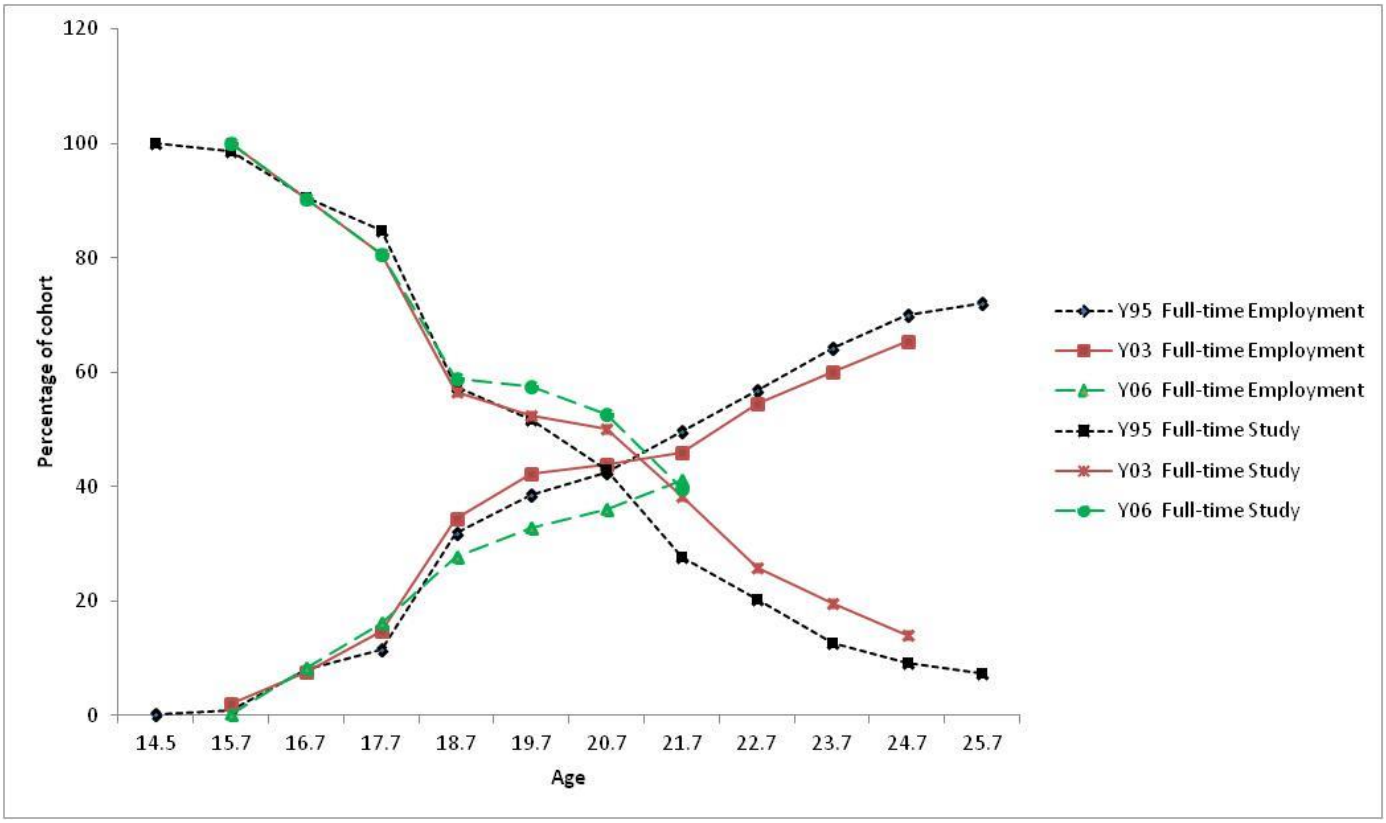


Figure 15 also makes it clear that more individuals are pursuing study: in the Y03 and Y06 cohorts the lines for full-time study are higher than those for the Y95 cohort. However, the level of full-time work during the period between 18 and 20 years of age is higher for the Y03 cohort than for Y95 and Y06, and the Y06 cohort has the lowest percentage of individuals in full-time work.[14]

Social factors

Bynner et al. (2002) show that young people are more likely to stay at home longer. For those who do move out, they are likely to be in a share house with their friends or to co-habit. Marriage has become less popular, and the age of women for their first birth has risen (ABS 2009).

Young men more than young women are more likely to remain at home (Cobb-Clark 2008). The age at which young people leave home has risen over time, with 19% of young people aged 20 to 34 living with their parents in 1986, compared with 25% in 2009 (ABS 2009). For young people aged 20 to 29, there was a 9% increase in the proportion of those living with their parents between 1976 and 2001 (ABS 2005).

During 2006—07, 31% of young people aged between 20 and 34 had left their parental home and had returned at some point (ABS 2009).[15]

Table 4 presents information from the Australian Bureau of Statistics (2009) that shows that in 2006—07 around 17% of people aged 25—29 still lived with their parents. Of this 17%, 9% had previously left home

[14] This could be attributed to the prevailing youth labour market conditions than a factor of the cohort itself.
[15] This includes 22% of those who were not living with their parents at the time of the survey.

and returned. Of the 83% of 25 to 29-year-olds who no longer live with their parents, 27% had left home and returned home at least once.

Table 4 People aged 20–34 years: moving out of, and back to, the parental home, 2006–07

	Age group (years)			
	20–24	**25–29**	**30–34**	**20–34**
	%	**%**	**%**	**%**
Total lives with parents	**47.2**	**16.8**	**8.2**	**24.5**
Has never left home	34.9	7.8	3	15.6
Left home and has returned	12.3	9	5.2	8.9
Total does not live with parents	**52.8**	**83.2**	**91.8**	**75.5**
Left home and has not returned	37.2	49.5	55.4	47.2
Left home and returned at least once	12.4	26.5	27.3	21.9
Has never left home, but lives separately from parents[(a)]	3.2	7.3	9.1	6.5
Total persons	**100.0**	**100.0**	**100.0**	**100.0**
	'000	'000	'000	'000
Total persons	**1 495.3**	**1 389.6**	**1 433.5**	**4 318.5**

Note: (a) Includes people whose parents may have died or moved away.
Source: Australian Bureau of Statistics, Australian social trends, June 2009, page 25, cat.no. 4102.0.

Figure 16 Percentage of cohorts living or not living at home

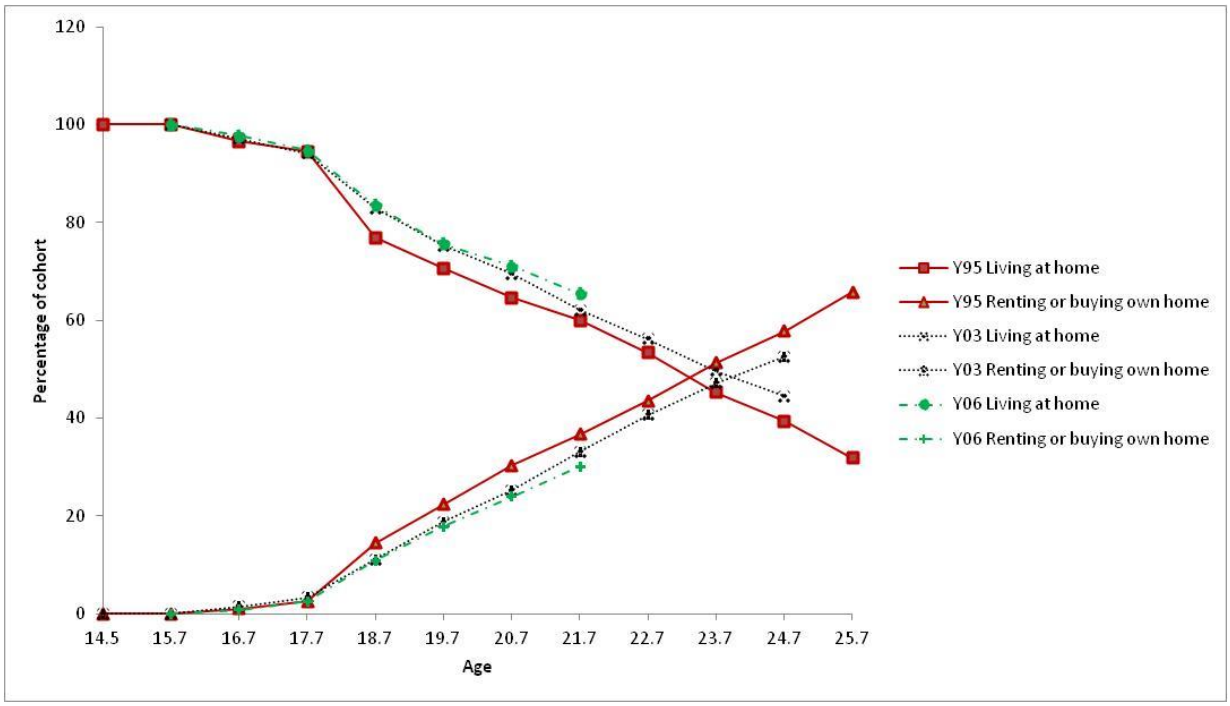


Figure 16 presents the proportion of young people from the Y95, Y03 and Y06 cohorts who report living at home or renting or buying their own home across ages. From this figure, it can be observed that the age at which the majority of young people who are buying or renting their own home is increasing, from around 22.7 years for the Y95 cohort, to around 23.7 years for the Y03 cohort. However, importantly, it can be seen that, even at age 25 years, the percentage of young people who live at home is greater than 30%.

On the whole, the evidence is that young people are remaining financially dependent on their parents for longer (see for example, Hartley 1993; Whittington & Peters 1996; Schneider 1999; Weston et al. 2001).

Advantages and disadvantages to extending LSAY

There are very real advantages to extending LSAY to beyond 25 years of age. Doing so would enable the transition points of all young people to become more visible. It would also allow for better estimation of the returns from education and training, particularly for those young people who pursue higher-level qualifications or who return to education in their 20s.

Extending LSAY would also enable better modelling of social outcomes, particularly investigating the living arrangements of those who leave home and then return to the parental home.

There are also clear disadvantages. There are additional costs associated with extending the length of time for sample maintenance, chasing contact details and conducting extra interviews. In terms of the last point, the cost of interviewing up until, say, age 30 could be offset by interviewing less frequently in some of the earlier waves (see table 5).

The greatest disadvantage to extending the survey could be on the influence on response rates and the impact on bias. It is likely that those who remain in the survey for this duration of time are those who are more advantaged, have higher achievement scores and are more successful in life.

Feasibility

It would be possible to maintain the contact details and to continue the sample maintenance with all existing LSAY respondents and it is worth highlighting that the Y03 cohort is currently in its planned final year of interviewing.

Extending the survey to, say, age 30 is practically feasible, and if considered, interviewing should be conducted less frequently than annually — perhaps every two years. If this approach were adopted, good sample maintenance would be necessary and contact lists would need to be maintained, ideally on the current six-monthly cycle.

Further, should extending the survey be considered, it would also be worth investigating the frequency of interviewing before age 25. It is important than annual interviews occur during the periods in which many significant transitions begin (ages 17 to 20). After age 20, it would be possible to move to a two-yearly interview cycle. This cycle would result in only the addition of a single interview wave. Table 5 presents an alternative interviewing cycle, which enables interviewing to age 30 with only one additional interview.

A disadvantage of a two-yearly interview cycle is respondents' recall error. This could be partially overcome through the use of life-course information during interviewing (that is, the respondents major life transitions and achievements are presented on screen to the interviewer). The fieldwork contractor would also have higher costs in rebuilding the sample from the previous wave as a greater number of sample members are likely to be lost in the two-year intervening period. However, the benefits in obtaining information for a longer period of time could justify this cost and effort. There is evidence of successful two-yearly interview cycles in other longitudinal surveys (Canadian Youth in Transition Survey, United States National Longitudinal Survey of Youth). A review of the types of research questions that older members are asked should also be undertaken for the later waves, and the questionnaire should be reviewed to better reflect the information required.

Table 5 Alternative interview schedule to extend LSAY to age 30

Age	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	No. waves
Existing interview schedule	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	10
Possible alternative schedule	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	11

Cost considerations

There will be an increase in costs to extend LSAY to age 30. However, savings could be made through the re-timing of the interviews, particularly as the cohort ages.

If interviewing became biennial from the time the cohort reaches the age of 20, then extending to age 30 would result only in an increase of one extra survey wave.

Extra costs would also be incurred as a result of the increased time required to maintain the samples. The contact database management and updates would need to be undertaken every year and, possibly, as happens now, every six months. Under the scenario of two years between interviewing, maintaining contact, relationships and communication with respondents would need to be sustained for a further five years.

Further consideration should also be given to rebuilding the sample from, say, age 20 for every wave thereafter. However, genuine hard refusals would need to be excluded. Increased costs in fieldwork may also arise, as it may be more difficult for the fieldwork contractor to locate individuals, although this may be partially offset through the sample maintenance efforts.

5. Health, social and emotional wellbeing, resilience and adaptability

Health

The health information, other than mental health, most likely to be sought in a national survey of adolescents and young adults relates to basic anthropometrics, physical activity and dietary behaviours.

Several surveys of Australian children seeking information on basic anthropometric data, nutritional intake and physical activity have been conducted in Australia: in 1985 (Department of Community Services and Health), 1995 (ABS), and 2007 (CSIRO & University of South Australia 2008).

In the 2007 Australian National Children's Nutrition and Physical Activity Survey (ANCNPAS), 4487 children aged from two to 16 years were surveyed. The survey used a combination of computer-assisted personal interview, conducted in the child's household, and computer-assisted telephone interview to collect data.

In the US, the National Longitudinal Study of Adolescent Health (Add Health) has been conducted with the first wave of data collection occurring in 1994—95 with 11to 18 year-olds (grades 7 to 12). Since then, a further three in-home personal interviews have been conducted (1995-1996, 2000-2001, and 2008). The oldest participants are now aged about 32 years (Harris et al. 2009).

Possible data items for a health module in LSAY

ANCNPAS provides a model for a health module in LSAY and includes data items under three main headings:

- food and nutrition
- physical and sedentary activities
- body size and shape.

In an ANCNPAS interview, recall of dietary intake and physical activity over three 24-hour periods was prompted. It is important to gather data for participants on different days, as dietary and physical activity regimes are likely to differ greatly between weekdays and weekends. This has implications for a possible health module in LSAY, where many questions about the two domains of diet and activity will need to be asked vis-à-vis school or work day and a weekend day.

Based on the recall of food and drink intake, it is possible to estimate energy consumption, total carbohydrates, sugars, starches, fibre, protein, fat (total, saturated and unsaturated), various vitamins and micronutrients. The coding of this information may require some input from dieticians as it may be difficult to create a look-up table to estimate nutrient intakes from the diverse list of foods likely to be identified by participants.

An alternative to an attempt to gather information on all foods and drinks consumed might be to seek information on specific food groups or particular foods within each of the major groups. This would enable typical energy and nutritional intakes to be estimated, which would simplify the data collection, although

it would result in the loss of detailed information about some nutrients. However, there may be advantages in targeting specific food intakes. These could include certain convenience foods, fruit and vegetables, and specific drinks, such as alcoholic beverages and their patterns of consumption.

To gather information on physical activity in the ANCNPAS, students were prompted to report on the two days prior to each of the two interviews. In an LSAY CATI, this could be difficult and it may be necessary to ask about specific selected activities, for example, the time spent walking, running, watching TV, working or playing on a computer, reading and resting. Membership of sporting and fitness clubs and frequency of use might also be sought. From these data, it would be possible to estimate the amount of time spent in sedentary, light, or moderate to vigorous physical activity. Such information would need to be sought for both school or work days and non-school or non-work days.

The third dimension of health data concerns physical data such as height and weight. While this information could be collected, its accuracy might be questioned. Gathering data on height and weight would enable a body-mass index (BMI) to be calculated. For broad surveys like LSAY, these two items might be adequate. These items were collected for the Y03 cohort, but have subsequently been dropped.

The Add Health survey collects data on participants' social and economic contexts and their psychological and physical well-being. The study gathers a variety of health status measures, and these are used as both predictor and outcome variables.

General health status

While it provides a useful model for a health module in LSAY, ANCNPAS does not include questions about general health status such as chronic medical conditions.

The current LSAY program includes an item in wave 3 on general health status: 'Do you have any disability or health problem which limits the amount or type of work you can do?' The purpose of this item is to elicit health status and its relationship to work.

LSAY is not an appropriate vehicle for collecting detailed and sensitive health data, but questions about common chronic conditions, for example, diabetes (currently included in LSAY), the use of common medications (analgesics), and the frequency of access to health services (visits to general practice clinics in the preceding year) could be included. In order to make a case for gathering any additional health status data, a relationship between them and the success of youth transitions would need to be demonstrated. The current general health status questions appear to be adequate for LSAY purposes.

No evidence has been found between the activity and diet measures reported in ANCNPAS and long-term health or other outcomes. This should not be expected, as the study was not designed as a longitudinal one. The long-term outcomes reported from the Add Health study have used health and wellbeing measures during adolescence as predictors of later health and well-being outcomes (Adam, et al. 2011; Hoyt, el al. 2012). The study has not been used to examine the sorts of outcomes that are of interest in a youth transitions study. Again, this is unsurprising, as Add Health is a health survey.

Feasibility and reliability of self-reported health data

While it is not feasible for a survey program like LSAY to capture the breadth and depth of data that specialist surveys do, ANCNPAS does include variables that could be collected in a health module in LSAY.

Data that could not feasibly be collected include objective measures of physical activity captured by pedometers or accelerometers. Other data that might be unreliable include some physical attributes such as waist measurement. In ANCNPAS, data such as these were collected in a personal interview. The

accuracy of self-reported weight and height data and recall of dietary and activity behaviours would need to be checked.

Data reliability is an important issue, particularly as much LSAY data is sought through self-reports. Participants in ANCNPAS were asked to recall aspects of dietary intake, and based on those reports, estimates of nutritional intakes were made. It was apparent in ANCNPAS, using BMI and claimed nutritional intakes, that some participants, particularly in the 14 to 16-year-old age group, under-reported their dietary intakes. It was estimated at between 8% and 16% of the 14 to 16-year-old age group under-reported their energy intake (CSIRO & University of South Australia 2008).

Advantages of a health module in LSAY

ANCNPAS includes data on children aged from two to 16 years. The sample sizes in each of the four age groups (2—5, 6—8, 9—13 and 14—16 years) targeted by the survey are modest and the response rate was 40%. The relatively low response rate could be attributed to the invasiveness of the survey, with an in-home personal interview being a requirement of the survey design. Moreover, the sample was not fully representative of the populations of children in the targeted age groups, as lower-income households appear to be under-represented. This is a weakness of the survey, since dietary and physical activity behaviours may be linked to social disadvantage.

The inclusion of a health module in LSAY would have the advantage that this module could be scheduled for inclusion in the survey in several waves, for example, waves 2, 4, 6, and 8, providing unique information on individual changes in, for example, BMI, diet and physical activity over time. This information would add to similar cross-sectional data that do exist. Moreover, the trajectories of diet and activity over time could be informative in a population in which obesity and numerous related chronic health risks and conditions are of growing concern, not least because of the projected long-term impacts on health budgets.

Social and emotional wellbeing, resilience and adaptability

While resilience and adaptability are identified as separate issues in the terms of reference for the investigation of possible topics for inclusion in the LSAY program, they are recognised as components of positive mental health and in this discussion are considered an aspect of social and emotional wellbeing.

The broad area of social and emotional wellbeing has been researched extensively over the past 20 years, with notable contributions from Seligman on positive psychology, including resilience, learned helplessness, depression, optimism and pessimism (for example, Seligman 1998); Goleman on social and emotional intelligence (Goleman 1995); and Salovey on emotional intelligence and the regulation of affect (Salovey, Hsee, & Mayer 2001). These developments have two key elements. They represent a move from viewing mental health as a set of problems that need to be cured, to viewing mental health as being a positive state. They also represent the recognition that positive mental health consists of a set of knowledge, skills and behaviour that can be taught and learned. This latter realisation has had a substantial influence on the role of schools in promoting social and emotional wellbeing.

There are good reasons for the increased interest in wellbeing. There is evidence that the incidence of depression and anxiety has increased, while life satisfaction has remained static, despite objective evidence of improving social circumstances (Levitt et al. 2007; Seligman et al. 2009). Higher levels of wellbeing are associated with improved academic learning, more successful transitions to adulthood and greater productivity (Dix et al. 2012; Levitt et al. 2007; Seligman et al. 2009).

Some high-profile schools in Australia, including Geelong Grammar School, have implemented the 'positive education' program promoted by Seligman. The KidsMatter program, developed for both primary and early childhood settings (<http://www.kidsmatter.edu.au>), includes four components, one of which is a social and emotional learning curriculum designed to enhance wellbeing. The KidsMatter program was evaluated and found to have made a statistically significant and practical difference to the lives of young children, with a reduction in reported negative behaviours (conduct problems) and an increase in mental health strengths (pro-social behaviours) (Slee et al. 2009).

The argument, therefore, for attending to wellbeing is compelling.

Assessing social and emotional wellbeing

The breadth of constructs associated with social and emotional wellbeing poses challenges to its measurement, especially in a program like LSAY.

Currently, the LSAY program includes a Life Satisfaction scale comprising 14 items. This set of items is included in all waves from wave 2 and it provides a trajectory of one aspect of personal wellbeing over time. Unpublished analyses of this scale show that it has good psychometric properties, but it is apparent that it does not capture some key dimensions of social and emotional wellbeing. Being a general life satisfaction scale, it does not measure the individual strengths that are elements of a broader conception of wellbeing. Life satisfaction might be an outcome that follows the effective deployment of positive social and emotional knowledge and skills.

Because of the breadth of wellbeing, many instruments are available to measure its components. These potential components include curiosity, creativity, life satisfaction, personal growth, psychological wellbeing, happiness and character strengths. A selection of potential instruments for measuring these and similar constructs is listed on the Positive Psychology Center website (<http://www.ppc.sas.upenn.edu/ppquestionnaires.htm>).

Levitt and colleagues (2007) reviewed many instruments used in the screening and diagnosis of children and adolescents for a range of mental health problems, from developmental delays in young children, to emotional and behavioural problems, to depression, anxiety and suicide among adolescents. For ethical reasons (see note below), it is important in the LSAY program to take a cautious approach to mental health screening and to focus on population-level screening instruments and to avoid those that are designed for targeted or at-risk groups, or those designed for use with already diagnosed clients. The review by Levitt et al. (2007) included instruments designed to be completed by teachers, parents and the individuals themselves. Their review considered the efficacy (accuracy), feasibility, and focus (narrow or broad) of instruments. Of the broad-spectrum instruments, the Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997) appears to meet the likely requirements of the LSAY program, as it can be self-completed by adolescents (aged 11—17 years) in 5—10 minutes.

Two major reviews of wellbeing measures have been undertaken in Australia. The ABS (2001) reviewed various wellbeing indicators, although its focus was on contextual factors, such as health, family and community, education and training, work, economic resources, housing, crime and justice, and culture and leisure, that influence successful outcomes. However, in the LSAY program, individual and direct measures of wellbeing are required.

In 2006 the Australian Institute of Health and Welfare (2012) commenced work on the measurement of headline indicators for aspects of children's health, development and wellbeing. In its 2012 report, the institute included a focus on wellbeing and identified two possible instruments, the Strengths and

Difficulties Questionnaire and the Australian Council for Educational Research (ACER) Social and Emotional Wellbeing Survey. The institute recommended SDQ.

While there is considerable support for the use of SDQ, there is little published evidence of its effectiveness in the Australian context. SDQ was used in the KidsMatter evaluations (Slee et al. 2009), although those evaluations used the teacher-completed versions, which is necessary with younger children.

In an unpublished review of SDQ in a study of wellbeing among secondary school students in South Australia using the self-report form, it was found that the instrument had some measurement deficiencies. The instrument, which has five sub-scales — emotional symptoms, conduct problems, hyperactivity, peer problems, and pro-social — was found to yield individual measures of relatively low precision and that some of the scales were poorly targeted for a general population. It is thought that the measurement difficulties are partly attributable to the limited response format used for the items (not true, somewhat true, or certainly true). An alternative instrument, the ACER Social and Emotional Wellbeing Survey, has better measurement properties, but takes much longer (about 30 minutes) to administer. Low measurement precision reduces the explanatory value in regression models of variables derived from the instrument.

In the LSAY program, it may be desirable to focus on certain aspects of the SDQ sub-scales, as even a five-minute instrument may be infeasible in an LSAY interview. Further, a revised version may be required: improving the precision of the individual measures derived from the instrument and extending the age range by altering the item wording so the same instrument could be used with adolescents and young adults would be desirable.

Lessons can be learnt from comparative longitudinal studies. The Millennium Cohort Study is a UK study that follows children who were born between 2000 and 2002. Hansen and Joshi (2007) note that children's emotional and behavioural problems were assessed using the SDQ. The SDQ was implemented as a computer-aided self-completion questionnaire, undertaken by parents. The data suggest that most children are emotionally adjusted and well behaved but the SDQ detects differences in behavioural problems between families from disadvantaged backgrounds compared with families from less disadvantaged backgrounds.

A key aim of the Millennium Cohort Study is to understand parental wellbeing and its influences on children. Data on the wellbeing of parents, including psychological morbidity, life satisfaction, drug use and problem drinking, were collected separately. For example, psychological distress was measured using the Kessler 6 scale, which is described as being a common instrument within general health surveys (Hansen & Joshi 2007). An evaluation of the impact of the Millennium Cohort Study on research, policy and practice by Johnson and Antill (2011) suggests that organisations are using the data on both parental and child wellbeing for a range of research questions.

The US National Longitudinal Survey of Youth 79 cohort measures attitudes and personality traits. As noted by Nguyen et al. (2010), self-esteem is measured using the Rosenberg self-esteem scale, which is a 10-item self report measure. Questions on delinquent and deviant behaviour are also included in the study. Questions on topics regarded as sensitive, such as on criminal activity and sexual behaviour, are administered via audio computer-assisted self-interview (ACASI) technology to improve response quality.

The Australian Youth in Focus study collects information on health, attitudes and the impact of life events via a separate self-completion questionnaire. Nguyen et al. (2010) note that the questionnaire is

administered online or through the mail to improve the quality of information. A monetary incentive of $15 is offered to improve response rates.

Data collection methods for health and social and emotional wellbeing

If self-reported health data such as height and weight and simple dietary and activity are gathered, the current CATI mode of administration is adequate. However, if additional health data are required, for example, if detailed diet and activity records are required, self-completed diaries or in-person interviews might be necessary. While diaries may be an attractive option, patterns of dietary and activity behaviours differ between weekdays and weekends and there may be quite different patterns for individuals. For this reason, in addition to sampling individuals, a sampling frame for days would need to be implemented. Respondent burden would need to be considered. This might mean asking participants to report on one weekday and one weekend day. Having respondents provide data on different days can impose some limitations on the data available for research purposes. The Time Use Survey (ABS 2008) is a difficult dataset to use because of these limitations.

In-person interviews are an attractive method for collecting detailed data. An obvious limitation is their cost — in training interviewers and in conducting interviews. It was necessary to use in-person in-home interviews for ANCNPAS but this seems to be associated with a high refusal rate.

The SDQ is a paper-based survey, although its format makes it amenable to CATI administration, as each scale has only eight items, each with three response options. It appears likely that it could be administered quite quickly, although it would need to be field-tested in the CATI administration mode.

If either or both the health and social and emotional wellbeing modules were added to the LSAY program, they would add to the time taken for interviews, and other sections would need to be reduced or removed to accommodate them. In relation to both health and social and emotional wellbeing, the inclusion of these modules at an early stage in the program (perhaps wave 2 or 3 if LSAY remains linked to PISA), and again at a later stage (perhaps wave 6 or later), would enable several research questions to be addressed. The early measures of health or social and emotional wellbeing might be useful predictors of their later status as well as of transition success.

One option for accommodating new measures such as health and social and emotional wellbeing would be to administer these modules to a subset of the sample, while other modules were administered to another subset of the sample. There is a precedent for this in LSAY in the trial of the social capital items administered to a sub-sample before a reduced set was administered later to all participants. This sub-sampling approach has the advantage that a wider range of measures can be collected. However, it has the disadvantage that the measures are not useful in large-scale models of transition success.

An alternative approach, especially for trialling modules, is to ask a sub-sample of participants to respond to a separate interview. Modest compensation may be offered for this additional burden. (See chapter 8 for further details on this approach.)

Ethics in measuring health and social and emotional wellbeing

The purpose of large-scale surveys like LSAY is to gather population-level data in order to inform policy development and to enhance practice. Although data are collected at the individual level, individuals are not normally the focus of the survey. In collecting data that could be regarded as sensitive or which might indicate that an individual is at some risk, it is incumbent of the agency responsible for the survey to take action to ameliorate any perceived risk (National Health and Medical Research Council, Australian

Research Council and Australian Vice-Chancellors' Committee 2009, pp.15—18). The survey itself is unlikely to create risk, but it may uncover individuals who are at risk, and therefore the responsible agency has an obligation to address that risk. This may be done by including, in material provided to participants, information about access to services that can assist participants (for example, counselling services), but it may require a greater level of intervention if the risk is deemed to be high. The development of protocols for the identification and management of risk should be considered.

Cost considerations in measuring health and social and emotional wellbeing

If self-reported health and wellbeing data are gathered by adding modules to the LSAY program, they would add to the time taken for interviews, and other sections would need to be reduced or removed to accommodate them. This would not result in additional fieldwork costs but there would be additional testing costs. Cognitive testing is estimated to cost at least $1000 per respondent. Due to the high costs, cognitive testing is usually only undertaken on a small sample. Pilot testing in CATI mode would be required on a larger scale than that of cognitive testing. Copyright costs will be incurred if using an established survey such as the SDQ. Questionnaire development costs will be incurred if not using an established survey.

Data collection costs will be higher if additional health data are collected through self-reported diaries or in-person interviews. It is estimated that in-person interviews can cost up to five times that of CATI interviews (Dillman, as cited in De Vaus 2002). However, the data collection approach and cost will need to be considered in conjunction with any other changes recommended by the review panel. For example, if a CATI interview is adopted as a data collection method for wave 1, it may be feasible to add health and wellbeing modules to this survey wave as the length of a CATI interview for wave 1 is estimated to be approximately 10 minutes. If in-person interviews are adopted for wave 1, it becomes more feasible and less costly to collect detailed information on diet and activity.

Summary

Instruments for the measurement of aspects of health and social and emotional wellbeing are available and could be adapted for use within the LSAY program. In the cases of both health and social and emotional wellbeing, clear policy advantages are expected from the measurement of these constructs in LSAY.

A clear purpose for these measures is required. Health and social and emotional wellbeing could be valuable progress measures in the transitions of youth into adulthood. If they are used as outcome indicators, they would need to be measured with some precision, and this has implications for the number of items required in the surveys. These constructs could be used as explanatory variables in further explorations of the outcome measures currently used in many LSAY analyses — successful completion of education and training and successful transitions into the labour force.

6. Administering a parent questionnaire

The lack of a parent questionnaire in the early stages of LSAY means that the scope and accuracy of information acquired may be limited. This chapter provides a discussion of the rationale for administering a parental questionnaire and the likely impact on analysis and reporting.

In the first wave of LSAY (PISA) students are asked about their parents' education. There is some missing data on these variables. For example, in PISA 2003 5% of mothers' and 8% of fathers' highest level of education are missing. Some of these 'missing' data are related to single-parent families. This is a problem because it is apparent from the cases where information is provided that low levels of parental educational attainment are associated with a higher likelihood of school non-completion and lower post-school participation in education and training. Further, the outcomes (for example, educational attainment) of students with missing data on home background variables are inferior to the outcomes of those for whom data are available. This suggests that some of the cases where data are missing are families where there is low parental education.

In addition to the problem of missing data on parental education is the accuracy of the data provided. In countries where the parent questionnaire was completed, students' and parents' responses to these questions can be compared (See table B3). For some categories of parental education, e.g. parents with university degrees, the agreement is over 80%, but for other categories, e.g. parents with a tertiary VET qualification, the agreement is below 50%. The likely impact of this level of inaccuracy requires some judgment. If the disagreement is systematic, e.g. if students consistently under- or over-state parental education, the problem is greater than if students randomly misreport parental education by one level.

Variation between students on outcomes such as achievement and educational attainment is overwhelmingly related to family circumstances, with a much smaller proportion of the variation being related to school factors. Typically, about 80% of the variation in outcomes, even those such as achievement that appear to be directly linked to in-school activity, are predicted by individual and family background factors and about 20% to school factors (see for example, Curtis & McMillan 2008; Gemici, Lim & Karmel 2013).

These two observations suggest a need to gather as much accurate home background data as possible.

Parental questionnaires may elicit information about parents and family circumstances about which students are unaware or about which they are unable to provide accurate responses. These issues are illustrated in the discussion that follows using data from the international studies, PISA and Progress in International Reading Literacy Study (PIRLS), where parental questionnaires were administered. In the PISA surveys, parental questionnaires are an option that countries may choose to administer. In 2009, 15 of the 65 participating countries elected to include the parent survey. In the 2006 PIRLS program, 40 countries participated and data are available for all but one of those countries.

The information sought in those surveys is indicative of the evidence of parental and family-related circumstances that could be elicited in order to understand the factors that influence youth transitions. However, in contemplating a parental survey, with the aim of enhancing the information obtained through the LSAY program, information other than that obtained from the PIRLS and PISA parent questionnaires could be considered. Those surveys, designed for cross-national administration, may not gather

information particularly relevant to the Australian context. Australia has similarities to Canada, historically and in its multicultural composition, but is unlike Singapore, which has three large cultural groups, and quite unlike Korea and Finland, which are substantially mono-cultural.

In the Y95 and Y98 LSAY cohorts, items were included in the student questionnaire that sought information about parental aspiration and expectation. These items, of course, elicited students' perceptions of their parents' aspirations, although those perceptions may not reflect parents' own views. Yet, there is evidence (for example, Campbell, Proctor & Sherrington 2009), that parental aspiration is very important in school choice and the support and encouragement that parents provide in the education of their children. Ethnicity was investigated as a factor in academic and social outcomes for youth in Australia by Marjoribanks (2002, 2004, 2005) using LSAY data, although he did not have access to parental aspiration. This has been the subject of investigation in the US and UK (see Spera, Wentzel & Matto 2009).

In the Y03 cohort, questions on social capital were trialled, initially with a sub-group of the cohort and later with the complete sample. Despite its great promise (Coleman 1988; Putnam 1993) and the considerable effort to define it (ABS 2004, 2006; Bullen 2004; Stone & Hughes 2002), social capital has not proven to be a useful predictor of transition success. This is disappointing from a policy perspective. By contrast, socioeconomic status does explain variation in many outcomes, for example, achievement, post-school education participation, and employment, but this is not amenable to direct interventions. That is, while it is possible to provide additional support to disadvantaged students, little can be done to change the circumstances that lead to the identification of disadvantage. The attraction of social capital is that aspects of it may be amenable to intervention. It is possible that better information on the social capital of families could be obtained from parents rather than their children.

In summary, consideration of including a parental questionnaire in LSAY would profit from an investigation of the instruments used in current international surveys like PIRLS and PISA, but the needs of Australian policy-makers must be considered and an investigation of a wider range of issues that might inform youth transition policy development should drive a possible parent questionnaire.

In the sections that follow, a brief summary of the issues that may arise in an attempt to implement a parent questionnaire are presented. This summary is organised as advantages and disadvantages of the initiative, its feasibility, the likely impact on the analysis and reporting of data, and the likely cost of the initiative.

Information derived from two large international surveys, PIRLS and PISA, is presented in appendix B, and that information is used in the discussion of the merits and challenges of the initiative.

Discussion of a parental questionnaire in LSAY

If the LSAY program remains linked to PISA, the PISA parent questionnaire could be administered. The PISA 2006 parent questionnaire included items about science and the environment; namely, their children's involvement in science activities (items that were in the student questionnaire), parents' views on the importance of science for future careers, and parents' views about the value of science and environmental issues to society. The questionnaire also included parents' perceptions of their child's school, the cost of education, their work, their educational attainment, and their family income. The science items were included because science was the major focus of the 2006 PISA program.

The logistics of administering a parent questionnaire need to be considered. If the questionnaire were administered through a CATI, it would be necessary to obtain a list of 15-year-old students from participating schools, to have parental contact details included, to select the students who it is hoped will

participate, and, for each student who does participate, to contact a parent to seek an interview. It is likely that the interview would be quite brief, probably taking less than eight minutes to administer.

If LSAY were separated from PISA, it would be feasible to use an instrument very similar to the PISA parent questionnaire. Approaches to parents could be made when the first wave of the survey was undertaken. The sampling method used in PISA (and in pre-PISA LSAY surveys) of stratified sampling of schools, followed by random sampling of students within schools, would mean that parents could be contacted using schools' parental contact details.

It may be possible, provided that parents consent, for schools to release the parental data they hold. However, the quality and consistency of information held by schools about parents would need to be checked.

Cost considerations

The main financial advantage in using an already developed parental questionnaire, such as the PISA parental questionnaire, is the absence of questionnaire development costs and a slight decline in costs associated with cognitive testing (estimated to be $1000 per respondent) as a smaller sample size would be required. Pilot testing costs would be incurred regardless of the questionnaire selected.

The approximate administrative cost to participate in the PISA parental questionnaire as a country add-on (based on 2012 costings) is less than $50 000. This includes domestic costs shared between the Commonwealth and the jurisdictions of just over $11 000 and international costs (Commonwealth cost) of at least $32 000. The international costs are based on securing participation by a minimum of ten countries. If more countries participate, the costs may decrease. It is not clear from the information available whether there would be additional data collection costs per participant, such as those outlined below, or if data collection costs are included in the PISA administrative costs.

The potential cost of data collection for the PISA parental questionnaire or an equivalent length parental questionnaire is examined for the following modes: personal interview one-on-one, telephone (computer-assisted telephone interview – CATI), web-based approaches and paper questionnaires. The data collection costs and the survey mode will need to be considered in conjunction with any other changes recommended by the review panel. For example, if an interview is adopted as a data collection method for wave 1, it would be feasible to sequentially interview both parents and students via the same interview mode and interviewer.

In 2009, the PISA parental questionnaire took approximately 20 minutes to complete (OECD 2013). CATI surveys of approximately 20 minutes can cost over $70 per participant when all costs, and not just the cost of the actual interview, are considered. The literature (for example, Doyle 2005; Groves et al., as cited in Weisberg 2005) and consultations with survey experts indicate that face-to-face interviews are at least double the cost of telephone interviews and could be up to five times the cost when all expenses are considered. Kaminska and Lynn (2013) add that costing face-to-face interviews is more complex when interviewing more than one individual in a household.

Face-to-face interviews can cost up to ten times that of web-based approaches (Kaminska & Lynn 2013; Richardson, Ampt & Meyburg 1995). Berrens et al. (as cited in Weisberg 2005, p.295) found that their 18-minute survey incurred similar costs for telephone interviews ($50 000) and web-based surveys ($60 000) for approximately 2000 completed questionnaires. They noted that the cost of the web-based mode increased to only $72 000 for 6000 completed surveys, showing that the cost of additional interviews is

low in the web-based mode. Web-based approaches tend to have high set-up costs but the data collection costs per participant are low once the infrastructure and survey is in place (OECD 2013; Weisberg 2005).

Howieson, Croxford and Howat (2008) add that web-based approaches are cheaper than paper-based surveys when the sample size is large. They estimate that the cost of printing and scanning for paper-based surveys exceeds the cost of programming and website hosting by over $60 000 for a sample size of 10 000. There is also a general consensus that paper-based surveys obtain lower response rates and result in poorer data quality compared with web-based approaches (Howieson, Croxford & Howat 2008; Richardson, Ampt & Meyburg 1995).

Advantages and disadvantages

There are clear advantages to a parental questionnaire. Such an instrument has the potential to elicit information that cannot be provided by students — or provided accurately by students. A parental questionnaire would be an opportunity to gather information not currently available in the student survey instrument.

There are very obvious disadvantages. Clearly, developing and administering the questionnaire will incur costs. Because models for a parent questionnaire are available, this cost is likely to be modest, although this will depend upon the constructs that are selected for inclusion. The greatest disadvantage is likely to be a modest response rate. Although the return rates of questionnaires vary considerably by country, those countries most like Australia (Canada and New Zealand) have rather modest return rates to PIRLS and PISA (see table B1 in attachment B). If a parent questionnaire is implemented, this potential barrier to success must be addressed, and a marketing program to encourage responses would be advantageous. A particular concern with modest response rates is the likelihood that the parents of students in more adverse circumstances are those least likely to respond.

Influence on analysis and reporting

Because there is a one-to-one correspondence between student and parental responses, the analysis and reporting of parental data is quite simple, since results from the parent questionnaire could be merged with the student-level data. Non-responses from parents may add a layer of complexity in the calculation of sample and attrition weights.

Analyses will be able to include data that is not currently available. This will add variables to analyses, but this should not add to the complexity of analyses. The use of additional variables in models will influence cross-cohort comparisons, as models developed with new variables (for example, parental aspirations) that exert a significant influence on outcomes (for example, school completion) may change the parameters that are reported for other established variables (for example, SES).

Summary

Whether LSAY continues to be linked to PISA or not, a parent questionnaire like the one used in PISA could be administered. The logistics of undertaking this are feasible.

There are potential advantages in gaining more accurate information about family backgrounds, especially parental education, occupations and incomes. The most likely disadvantage is a low parental response rate that would limit the value of the information collected.

7. Data linkage with LSAY

No single data source in Australia currently provides longitudinal data on young people's developmental trajectories from early childhood up to tertiary education and entry into the labour market. LSAY collects detailed information on young people's background characteristics, educational achievement and key life events. Administrative collections such as Medicare Australia and Centrelink or point-in-time collections such as the Australian Census also contain important data on factors that directly or indirectly influence young people's transition outcomes. Combining elements of different data sources can potentially generate a coherent data stream that cannot otherwise be gained from a single survey or administrative collection.

What is data linkage?

Data linkage refers to the process of matching records held in different data sources about the same person (Jutte, Roos & Brownell 2011). It applies to situations in which the different data sources to be combined contain information on the same individuals.

Significant work based on linked life-course data has recently been undertaken in Canada. Using linked data from the Manitoba Population Health Research Data Repository,[16] researchers have established important causal relationships between early life risk factors and long-term health, education and labour market outcomes (Brownell et al. 2010; Jutte et al. 2010; Jutte, Roos & Brownell 2010). Similar cross-sectoral studies with linked data have been conducted in Sweden (Lawlor et al. 2006; Li, Sundquist & Sundquist 2010).

Current Australian examples of cross-sectoral linkage projects include the prediction of reading and numeracy skills from early childhood development data (Gregory 2012) and the impact of social and clinical background factors on school readiness (Lynch 2012). A brief summary of selected research studies using linked data is provided in table 6.

[16] The Manitoba Population Health Research Data Repository links the population registry to several health, education and welfare databases. For details see University of Manitoba (2012).

Table 6 Summary of selected research studies using linked data

Study focus	Author(s)	Location	Data sources
Associations of child socioeconomic status and mortality	Lawlor et al. (2006)	Sweden	• Swedish Multi-generation Register • Swedish Cause of Death Register • Swedish Census
Linking Millennium Cohort data to birth registration and hospital episode records	Hockley et al. (2008)	UK	• Millennium Cohort Study • Birth registration data • Hospital record data
Impact of adolescent socioeconomic status on higher education participation	Chowdry et al. (2008)	UK	• English National Pupil Database • Higher Education Statistics Agency Student Records
Academic and social outcomes for high-risk youth	Brownell et al. (2010)	Canada	• Manitoba Population Health Research Data Repository
Biologic versus social predictors of childhood health and educational outcomes	Jutte et al. (2010)	Canada	• Manitoba Population Health Research Data Repository
Social, educational and medical outcomes for children of teenage mothers	Jutte, Roos & Brownell (2010)	Canada	• Manitoba Population Health Research Data Repository
Effects of parental occupation on low birth weight	Li, Sundquist & Sundquist (2010)	Sweden	• WomMed II[17] • Swedish Census
Development characteristics at age five as predictors of reading and numeracy skills three to seven years later	Gregory (2012)	Australia	• Australian Early Development Index • Western Australian Literacy and Numeracy Assessment • NAPLAN
Impact of social and clinical background factors on school readiness	Lynch (2012)	Australia	• Births, Deaths and Marriages • Children, Youth and Women's Health Service: Child Health Record (0–4 years) • SA Health: Integrated South Australian Activity Collection, Emergency Department Data Collection, Perinatal Data Collection • SA School Enrolment Census, NAPLAN, Australian Early Development Index
Does scored VET in Schools help or hinder access to higher education in Victoria?	Polidano, Tabasso & Zhang (forthcoming)	Australia	• NAPLAN data • University entrance data • University preferences data
Electronic health record linkage in the Millennium Cohort Study	Dezateux et al. (work in progress)	UK	• Millennium Cohort Study • Electronic health record of hospital admissions • General practice records

Note: Studies are listed in chronological order. Full citations are provided in the references section.

[17] WomMed II is a nationwide database that contains information from the Swedish medical birth register, which includes both birth records and prenatal care data (Swedish Centre for Epidemiology 2003). The WomMed II database also contains nationwide individual-level hospital diagnoses and death register data, as well as census data.

The Longitudinal Study of Australian Children (LSAC) provides a successful model of linking administrative data to a flagship longitudinal survey. Major efforts have been undertaken to link this study to the following administrative datasets (Soloff et al. 2007):

- health and development information recorded in parent-held records about every child after birth
- hospital records of the child's birth
- immunisation records held by the Australian Childhood Immunisation Register
- episodes and types of healthcare utilisation funded by Medicare Australia
- information held by Medicare Australia on the Pharmaceutical Benefits Scheme
- data on the quality of childcare centres and family day care schemes, as held by the National Childcare Accreditation Council
- ABS indices of disadvantage, remoteness indicators and other measures of interest (such as unemployment rates) derived from census data.

Further to the administrative collections listed above, Daraganova, Edwards and Sipthorp (2013) recently illustrated the process of linking NAPLAN academic achievement scores to corresponding LSAC participants. The link between the NAPLAN and the Longitudinal Study of Australian Children allows researchers to determine the impact of individual and parental background characteristics, early childhood and school interventions, and personal attitudes and aspirations on academic outcomes in Years 3, 5, 7 and 9.

A more detailed description of research using data linkage is provided in a recent discussion paper by Gemici and Nguyen (2013).

Advantages of data linkage with LSAY

The key benefits of linking LSAY to existing administrative collections are:

- Linked administrative data from the education, training, and health sectors would allow researchers to explore key drivers of young people's transition outcomes over a much longer time span than is currently possible through LSAY alone. In particular, this means extending LSAY with information on outcomes prior to age 15 (when LSAY starts) and beyond age 25 (when LSAY ends).
- Data linkage can significantly broaden the informational value of LSAY without adding to respondent burden. The benefits are particularly strong in topic areas that are currently limited in LSAY, such as health information and early educational performance.
- Data linkage can improve the accuracy of LSAY data because currently data are mostly self-reported and thus dependent on respondent recall.
- Data linkage would allow scope for adding new questions to LSAY because certain questions that are currently in the survey would be dealt with through linked data sources. New LSAY questions could then cover areas such as social capital and wellbeing in more depth, as well as additional measures of personal characteristics such as the personality traits and aberrant behaviour associated with young people's decision-making and impacts on later outcomes. Finally, a series of reflection questions could be added in later waves about influential events that impacted on their chosen transition pathway.

- Data linkage would greatly enhance the value of NAPLAN data by complementing student scores with a richness of individual background information from LSAY. This would add unparalleled value to NAPLAN data.

A more detailed description of these advantages is provided in a recent discussion paper by Gemici and Nguyen (2013).

Challenges of data linkage with LSAY

Legal challenges

The principal legal challenge is that of protecting privacy and data confidentiality, which includes obtaining written consent and developing specific protocols for the secure management of linked data.

Firstly, any potential data linkage between LSAY and administrative collections will require obtaining consent from respondents to link their data. Because there is no minimum age at which an individual can make decisions on his or her personal information under the *Privacy Act 1988*,[18] LSAY respondents can theoretically be asked directly for their consent.[19] However, it might be safest to seek written parental consent in addition to respondent consent until the LSAY respondent reaches the age of 18. In practice, this could be accomplished via mailing information packs to participants and their parents. These packs would contain consent forms in addition to various other communication materials. (For details, see the communication strategies in chapter 3 on attrition.)

Secondly, the development of a data-linkage process using LSAY data needs to take into account specific protocols that protect the privacy and confidentiality of the data and manage the safety and security of linked data (National Statistical Service 2010). These include the requirement of de-identifying linked data, use of an independent agency as data custodian, and secure storage of linked data. For data-linkage projects involving Commonwealth data for statistical and research purposes, an official 'integrating authority' must be used. An integrating authority is the single agency ultimately accountable for the implementation of a statistical data-linkage project (National Statistical Service 2012). Currently, the two integrating authorities for Commonwealth data are the Australian Bureau of Statistics and the Australian Institute of Health and Welfare (AIHW). This means that the actual process of linking any Commonwealth administrative collections to LSAY has to be coordinated by either one of these authorised external intermediaries. One important exception is the use of NAPLAN data for linkage with LSAY because NAPLAN scores are owned and held by individual state/territory governments.

[18] A general principle in determining when a young person has the capacity to make a decision on his or her behalf is when they have sufficient understanding and maturity to comprehend what is being proposed. In some circumstances it may be appropriate for a parent or guardian to consent on behalf of a young person where the child is very young or lacks the maturity of understanding to do so themselves (Australian Law Reform Commission 2008).

[19] Note that other guidelines adhered to by collection agencies or contractors may set a minimum age requirement. For example, the Australian Market and Social Research Society's (AMSRS) Code of Professional Behaviour, which is adhered to by the current fieldwork contractor for LSAY, requires researchers to obtain consent from a parent or guardian before a child of 14 years and under can be interviewed. Moreover, some ethics committees may suggest that whether consent from a 15-year-old is acceptable depends on whether it can be reasonably argued that they understand fully what they are consenting to. Best practice would likely seek parental consent in addition to respondent consent until the LSAY respondent reaches the age of 18.

Technical challenges

From a technical viewpoint it is feasible to link administrative data with LSAY because respondents' representative identifiers (that is, names, addresses, dates of birth etc.) are known to the LSAY data collection contractor (currently Wallis Consulting Group). However, certain technical challenges do exist.

Consent bias

As discussed above, legal consent needs to be sought from LSAY respondents or their parents/legal guardians in order to use LSAY data for linkage projects. Therein lies the challenge of consent bias. Consent bias denotes the fact that population groups with certain socio-demographic characteristics (for example, low-SES and Indigenous youth, as well as other equity groups) have lower consent probabilities for data linkage (Australian Institute of Health and Welfare & ABS 2012; Kho et al. 2009; Knies, Burton & Sala 2012; Sala et al. 2012). This can lead to an under-representation of already vulnerable population groups in statistical analyses with linked data. In LSAY, consent bias could be addressed to some extent by developing appropriate statistical weights based on consent probabilities.

Population overlap

Data linkage applies to situations in which the different data sources to be combined contain, at least in part, information on the same individuals. For instance, it is feasible to combine records from Medicare Australia with LSAY via data linkage because all LSAY respondents should, in theory, be included in the Medicare database.[20] Data linkage becomes less useful as the population overlap between data sources decreases, and the method is not applicable when attempting to combine data sources that do not contain information about the same individuals.

Technical and practical feasibility

There are numerous administrative collections held at different levels of government that could enhance the breadth of LSAY through data linkage. Datasets that collect information in areas where LSAY is limited are a natural starting point: information on outcomes prior to age 15 (when LSAY starts), information beyond age 25 (when LSAY ends), and health information.

A recent discussion paper (Gemici & Nguyen 2013) explored the feasibility of linking LSAY data to a number of different administrative data collections, including NAPLAN, Medicare, Centrelink, the Australian Census, as well as the Higher Education Statistics collection and the National VET Provider collection. One conclusion from the paper was that NAPLAN and Medicare would initially be the most feasible candidates for data linkage with LSAY. An overview of advantages and limitations of linking LSAY to NAPLAN is provided below.

Linking LSAY with NAPLAN

NAPLAN was introduced in 2008 to assess all students in Years 3, 5, 7 and 9 in reading, writing, language conventions (spelling, grammar and punctuation) and numeracy. Since 2003, LSAY has also featured cognitive assessment data for 15-year-olds in reading, mathematics, and science literacy through its

[20] LSAY participants would have Medicare records either on their own account or through their parents or legal guardians. The availability of Medicare records might be limited for recent immigrants to Australia.

integration with PISA.[21] The benefit from linking NAPLAN scores to LSAY would be access to literacy and numeracy development from Years 3 to 9, allowing researchers to control for academic achievement at earlier ages. Conversely, the lack of contextual information in NAPLAN data can be broadened with individual background and transition data collected from LSAY.

Unlike census data, which are held and managed centrally by the ABS, NAPLAN scores are stored by individual state/territory governments. At the same time, the Australian Curriculum, Assessment and Reporting Authority is a central repository of NAPLAN data from all states and territories. Against this backdrop, two options exist for approaching a data-linkage project between LSAY and NAPLAN: one option is to collaborate with each of the eight state/territory education departments separately; the other option is to work in collaboration with ACARA.

The advantage of collaborating with the Australian Curriculum, Assessment and Reporting Authority is that of having to interact with only one principal project partner rather than eight different state/territory partners. This is likely to greatly improve timeframes for an LSAY—NAPLAN linkage project. However, the disadvantage of collaborating with ACARA is that their combined NAPLAN file does not contain information on students' name and date of birth, which complicates the linkage process and may decrease its reliability to some extent.

The steps involved in collaborating with the eight state/territory education departments are as follows:[22]

1. Approach the head of the education department in each jurisdiction in order to introduce the project and outline the rationale for requesting access to that jurisdiction's NAPLAN data for linkage with LSAY. Also, list the wider benefits and steps involved in the process.
2. Obtain written consent for data linkage from LSAY respondents who are 18+ years of age, or from parents/legal guardians for participants who are minors. [23]
3. Engage a data-integrating authority such as the ABS or the Australian Institute of Health and Welfare. This will ensure that the linkage process meets the highest technical and privacy/confidentiality standards.
4. Wallis Consulting sends a complete data file of LSAY respondents to the data-integrating authority. This file also contains representative identifiers (first name, last name, school and date of birth) for each respondent.[24]
5. The jurisdictions send the NAPLAN scores and representative identifiers (first name, last name, school and date of birth) to the data integrating authority.
6. At this point, the data-integrating authority has two files: one complete data file of consenting LSAY respondents, including representative identifiers; and one data file, which contains only the NAPLAN scores and representative identifiers. The data-integrating authority matches the two files and then deletes the representative identifiers from the matched file.

[21] Prior to the link between LSAY and PISA from 2003 onwards, the Y95 and Y98 cohorts featured reading and numeracy tests that were administered to students in the respective base year. Test results led to the creation of three school achievement measures: achievement in literacy, achievement in numeracy, and combined achievement in literacy and numeracy. Further details are provided in NCVER (2011).

[22] NAPLAN scores have been successfully linked to the Longitudinal Survey of Australian Children (see Daraganova, Edwards & Sipthorp 2013). The steps that were undertaken to carry out the link between LSAC and NAPLAN are similar to those described here for a potential link between LSAY and NAPLAN.

[23] LSAY respondents' representative identifiers are held by the LSAY data collection contractor (currently Wallis Consulting).

[24] In LSAC, school and postcode were used as additional identifiers for the matching process.

7. The data-integrating authority returns the matched file to NCVER. This file now contains all of the original LSAY data as well as unit record NAPLAN scores. However, it does not contain any representative identifiers.
8. NCVER verifies the matched file and makes it publicly available via the Australian Data Archive.

The steps involved in collaborating with the Australian Curriculum, Assessment and Reporting Authority are similar except for the first step. Instead of approaching all eight jurisdictions, approval for data linkage would have to be sought only from the ACARA governing board.

Cost considerations

Based on previous Australian examples of data linkage,[25] the process of obtaining NAPLAN data and linking LSAY data with NAPLAN data could take at least six months for the first data linkage attempt. The cost to undertake a pilot project is therefore likely to be over $50 000, noting that a data linkage demonstration project with LSAY and NAPLAN data is a component of the 2013–14 LSAY research and analysis program. The time taken to engage with the owners of the NAPLAN data and undertake the linkage thereafter would be reduced as a process to obtain consent, obtain the NAPLAN data, and link the datasets would be established.

The costs will vary for the LSAY and NAPLAN data linkage demonstration project and future data linkage exercises depending on the approach taken to obtain consent from participants, obtain NAPLAN data and undertake the data linkage. The costs are likely to be lower if data consent can be obtained from participants during the LSAY fieldwork or as an add-on to a communication pack, rather than through a separate mail-out or CATI process. As mentioned, it is best-practice to obtain written consent for data linkage. The cost of obtaining written consent is much higher than administrating an online consent form when there are a large number of participants. The costs will therefore fall if it is established that it is not a legal requirement to obtain written consent.

The administrative and project management costs are likely to be lower if NAPLAN data are obtained from one organisation, such as the Australian Curriculum, Assessment and Reporting Authority, rather than each jurisdiction. The costs may also vary depending on the data-integration authority used for the linkage process.

Summary

Understanding youth transitions requires information on young people's individual background characteristics and the circumstances under which they grow up. Such information includes family and community background, physical health, psycho-social development, as well as academic achievement and the broader school environment. The ability to assemble this information into a coherent data stream from infancy right through to adulthood is invaluable for developing effective policy settings. In addition to informing policy-makers and practitioners about the need for policy intervention, such comprehensive life-course data can shed light on the question of when different interventions have an impact on transition outcomes.

Data linkage can provide a means to understand issues across the life course, without increasing respondent burden. Linking LSAY with NAPLAN data was recommended by Gemici and Nguyen (2013) as a

[25] This is based on the experiences of Daraganova, Edwards and Sipthorp (2013) and Polidano, Tabasso and Zhang (forthcoming).

first step to demonstrate the data-linkage potential. This linkage exercise also provides a means to explore correlations between NAPLAN and PISA, which is of direct relevance to the question of whether to maintain the current LSAY and PISA link, discussed in an earlier chapter of this report. The recommended approach to linking with NAPLAN data is based on the model used by Daraganova, Edwards and Sipthorp (2013) to link NAPLAN and Longitudinal Survey of Australian Children data. The results of this NAPLAN and LSAC data-linkage project show that technically it is possible to link survey data with NAPLAN data and there is reason for doing so. Privacy may be a key concern to the stakeholders involved, but using a data-integrating authority should alleviate any concerns held by the state and territory education departments who own and hold the NAPLAN data.

8. Conducting topical surveys, interviews and focus groups

Supplementary surveys, interviews and focus groups can be used to further examine specific youth transition issues. For example, an upcoming report (Halliday Wynes and Nguyen 2014) involved conducting interviews with the 2006 LSAY cohort in order to examine the impact of financial stress on young people undertaking tertiary study. This study combined quantitative analyses of the 2006 LSAY cohort, based on data already collected in LSAY (including data on government payments, income, and credit card activity) and qualitative data collected through interviews with a subset of participants.

Advantages and disadvantages

Supplementary surveys, interviews and focus groups provide key benefits compared with attaching additional modules to the questionnaire to improve questionnaire content as they:

- provide timely data on a specific area of interest to policy-makers compared with the time lags associated with adding new questions and modules to the main questionnaire through questionnaire development, fieldwork and data preparation stages.
- significantly broaden the informational value of LSAY without significantly adding to respondent burden or impacting questionnaire length. The benefits are particularly strong in topic areas that are currently limited in LSAY, such as health and wellbeing information.
- can be used on a smaller subset of participants whereby it is more feasible to offer incentives to encourage participation.

Technical and practical feasibility

The approach by Halliday Wynes and Nguyen (2014) provides insight into the likely costs and steps involved in conducting one-off surveys, interviews and focus groups. In the Halliday Wynes and Nguyen study, NCVER commissioned the Wallis Consulting Group to recruit and interview 50 respondents (by phone[26]) who were experiencing financial stress and who were currently enrolled in tertiary education. This Wallis Consulting Group was used for recruitment and interviewing for three reasons. The first reason was to maintain and build on the relationships between the current fieldwork contractors and the respondents. The second reason was to capitalise on Wallis's experience at recruiting and interviewing respondents. Finally, respondent information is held by Wallis and may only be passed to a third party once consent is given by participants.

The Halliday Wynes and Nguyen (2014) study was conducted within the LSAY research and analysis program, funded by the Commonwealth Department of Education, and with advice and support from the LSAY management team. Data access protocols may need to be revised to assess the value and impact of requests from external researchers to conduct further studies on LSAY participants. Key considerations would be whether increased burden placed on respondents could impact respondents' future participation in the program and the protection of individuals' privacy.

[26] 11 of the 51 respondents completed the study via an online mode. The remainder of participants were interviewed via the phone.

Cost considerations

In the study by Halliday Wynes and Nguyen, interview question development, recruitment and interviewing stages (based on a 30 minute telephone interview) were costed at approximately $25 000, with an additional $2500 included for a $50 incentive to encourage participation. This equates to a total cost of over $500 per participant. Costs could be minimised by using focus groups but this may reduce the quality of responses as some respondents may not be comfortable discussing personal circumstances with others present (Richardson, Ampt & Meyburg 1995). If interviewing a large number of participants, the cost could be reduced by using a web-based approach whereby the fieldwork cost per participant is very low once the survey has been developed and set-up (OECD 2013; Weisberg 2005).

Summary

Previous chapters of this paper (chapters 4, 5 and 6) have shown that the value of LSAY to policy-makers could be enhanced by improving the information available on outcomes beyond age 25 years, health and wellbeing, and parental background. There are also emerging issues relating to youth transitions that would be useful to policy-makers that are not specifically captured in LSAY. The main barriers to collecting this information are related to costs and respondent burden. Conducting supplementary surveys, interviews and focus groups can be used as a less costly alternative to altering the questionnaire content. The study by Halliday Wynes and Nguyen (2014) demonstrates that this approach can provide insights into a specific issue without adding significant costs by conducting studies on a smaller subset of the LSAY respondents.

Conclusion

It is clear that LSAY's ability to track young people over time makes it a valuable dataset for researchers in examining relationships between factors that impact on youth transitions. This key feature of LSAY, and longitudinal surveys in general, enables policy-makers to make decisions based on robust evidence. Indeed, our analysis of LSAY data and the views of best practice from the research confirm that LSAY provides reliable estimates of Australian youth, including when analysed at the jurisdictional and school sector levels.

As noted in the earlier stocktake report by Nguyen et al. (2010), there are a number of limitations to LSAY[27], which could be overcome by enhancing the survey in several areas. Our assessment confirms these limitations. LSAY suffers from attrition, particularly from the first wave (PISA) to the second wave. Further, the data suggests the current level of attrition impacts on the accuracy of estimates of important sub-populations. LSAY is also limited in its ability to provide in-depth analyses of health and wellbeing and capture accurate information on parental background. Research suggests that youth transitions are taking longer, providing a clear argument in favour of extending the survey beyond 25 years.

This paper examined eight enhancements to LSAY with the potential to improve its value and usefulness — without detracting from the real strength of the survey in investigating changes to the cohort over time. The enhancements considered in this paper (including changes to the sample design, addressing attrition, linking to other datasets, and improving questionnaire content) all have merit. They are also interrelated, in that an improvement in one area is likely to positively or negatively affect other aspects of the survey.

Improving questionnaire content by enhancing information on health[28] and wellbeing and parental backgrounds and extending the survey beyond age 25 years will improve the ability of researchers to analyse youth transitions. However, these options will incur costs, and the accuracy of the information will be dependent on response rates. Due to these common disadvantages, the first considerations should be attrition and the tweaking of the sample design to allow for cost savings and the reallocation of funds to enable other future improvements, such as a parental questionnaire.

The options for altering the sample design and the strategies for reducing attrition are somewhat dependent on whether the sample continues to be selected from school students who participate in PISA. However, even if the link with PISA is maintained, there are still options to alter the sample size and design, reduce the frequency with which new cohorts commence, address attrition and to alter the within cohort interviewing schedule. Table 7 provides a summary of the enhancements according to whether or not the PISA—LSAY link is maintained. As noted in the table, the removal of the link provides more options for changing the sample design. On the other hand, this approach will incur great costs, recruitment challenges and information gaps relating to schools and academic performance. Data linkage will be critical to determining whether NAPLAN offers a reliable alternative to the academic performance measures in PISA.

[27] These limitations are not necessarily unique to LSAY — they are weaknesses of longitudinal surveys in general.

[28] There is less evidence that improved information on diet and activity would improve research on youth transitions, compared with enhanced information on wellbeing and outcomes beyond age 25.

Table 7 A summary of LSAY enhancements by whether the link to PISA is maintained

Enhancements	Without PISA–LSAY link	With PISA–LSAY link
Frequency of starting cohorts	Any change to the frequency of starting cohorts is possible (including following one cohort over time) and would provide resources to reinvest into other enhancements.	There is an option to extend the time between cohorts to provide resources to reinvest into other enhancements. Options include every three, six and nine years or follow one cohort over time.
LSAY sample size	There are more options to oversample equity groups. This option could impact on the representativeness of the general Australian population.	The Indigenous group is the only sub-population that can be oversampled. This already occurs.
	The sampling could be modified to improve the ability to analyse sub-populations of schools: • improving the representative populations of schools in areas such as socioeconomic status • surveying a smaller number of students from a greater number of schools.	There are limited options to modify sampling to improve the ability to analyse sub-populations of schools.
	Reduce the sample size and reinvest resources into reducing attrition. This option would impact on the ability to analyse important sub-populations using LSAY, unless resources were reinvested into improving attrition.	Tweak the sample size (minimum under PISA is 5000) and reinvest resources into reducing attrition. This option would impact on the ability to analyse important sub-populations using LSAY, unless resources were reinvested into improving attrition.
Extending the survey to beyond age 25	PISA does not alter the options. Costs can be reduced by adopting a two-yearly interviewing cycle after age 20. Note, an alternative option is data linkage or running an ad-hoc survey.	PISA does not alter the options. Costs can be reduced by adopting a two-yearly interviewing cycle after age 20. Note, an alternative option is data linkage or running an ad-hoc survey.
Attrition	Options to address attrition are not limited. Options include offering incentives, communicating with participants and their parents (including on commencement of the survey), altering the survey mode, interviewing participants in wave 1 and modifying the questionnaires to make them more interesting.	Invest resources into maintaining sub-populations of interest or re-sample/rebuild the sample. Other options to address attrition may be limited by opportunities to alter the PISA data collection process in the first year. Addressing attrition from the PISA wave to the second wave will be critical. This could occur through offering incentives, interviewing participants in wave 1 and providing financial compensation to schools to provide accurate contact details.
Data linkage	Consent for data linkage can be obtained at commencement of the survey.	Options to obtain consent under PISA could include sending a consent form in a communication pack to participants, which may need to occur in wave 2.
Administering a parent questionnaire	The development and testing of a parental survey will incur costs. The development and testing costs may be slightly reduced if using models already available, such as the PISA parental questionnaire. Response rates will be critical with or without PISA.	The development and testing of a parental survey will incur costs. The development and testing costs may be reduced if using models already available, such as the PISA parental questionnaire. Response rates will be critical with or without PISA.
Improving information on health and wellbeing	The mode of delivery, survey instrument and the impact on the survey length will be critical with or without PISA. An alternative option is to include the topic in an ad-hoc survey or study or to link LSAY to a relevant dataset.	No difference, aside from there being no option to collect additional data in the PISA wave. The mode of delivery, survey instrument and the impact on the survey length will be critical under both options. An alternative option is to include the topic in an ad-hoc survey or study or to link LSAY to a relevant dataset.

Data linkage and the introduction of supplementary surveys and qualitative studies are highlighted as areas in which LSAY could be further enhanced without significantly adding to response burden and with minimal costs. In particular, the prospect of linking LSAY to existing administrative datasets is potentially the most promising undertaking in the years to come. Such data-linkage and supplementary projects have the potential to ameliorate current information gaps as well as address emerging areas of interest to policy-makers.

As noted, there are many options for enhancing the value of LSAY to policy-makers, practitioners and researchers. In prioritising the areas of most value, it will be important to be mindful of the strengths of LSAY. Hence, some options would improve the value of LSAY, such as oversampling sub-populations, but may introduce sample bias and thus reduce the accuracy of the sample at the Australian and jurisdictional levels.

References

Adam, EK, Chyu, L, Hoyt, LT, Doane, LD, Boisjoly, J, Duncan, GJ, Chase-Lansdale, PL, & McDade, TW 2011, 'Adverse adolescent relationship histories and young adult health: Cumulative effects of loneliness, low parental support, relationship instability, intimate partner violence, and loss', *Journal of Adolescent Health*, vol.49, pp.278—286.

Aliaga, A & Ren, R 2006, 'Optimal sample sizes from two-stage cluster sampling in demographic and health surveys', *Demographic and Health Research (DHS) Working Papers*, no.30, United States Agency for International Development, viewed 14 October 2013, <http://www.measuredhs.com/publications/publication-WP30-Working-Papers.cfm>.

Antonovics, K, Haveman, R, Holden, K & Wolfe, B 2000, 'Attrition in the new beneficiary survey and follow-up and its correlates: polling data: statistical data included', *Social Security Bulletin,* Spring, vol.63, no.1, pp.40—52.

ABS (Australian Bureau of Statistics) 2001, *Wellbeing*, cat.no.41600, ABS, Canberra.

——2004, *Measuring social capital: An Australian framework and indicators*, cat.no.1378.0, ABS, Canberra.

——2005 *Australian Social Trends*, cat.no.4102.0, ABS, Canberra.

——2006, *Aspects of social capital*, cat.no.4911.0, ABS, Canberra.

——2008, *How Australians use their time 2006*, cat.no.4153.0, ABS, Canberra.

——2009, *Australian social trends*, cat.no.4102.0, ABS, Canberra.

——2013, *Schools, Australia, 2012*, cat.no.4221.0, ABS, Canberra.

Ainley, J, Malley, J & Lamb, S 1997, *Thematic review of the transition from initial education to working life: Australia: background paper*, Australian Council for Education Research (ACER), Melbourne.

Australian Curriculum, Assessment and Reporting Authority (ACARA), My School website, viewed October 2013, <myschool.edu.au>.

AIHW (Australian Institute of Health and Welfare) 2012, *Social and emotional wellbeing: Development of a children's headline indicator: Information paper*, AIHW, Canberra.

AIHW & ABS 2012, *National best practice guidelines for data linkage activities relating to Aboriginal and Torres Strait Islander people*, AIHW, Canberra.

Australian Law Reform Commission (ALRC) 2008, *For your information: Australian privacy law and practice*, ALRC, Canberra.

Berk, ML, Mathiowetz, NA, Ward, EP & White, AA 1987, 'The effect of prepaid and promised incentives: results of a controlled experiment', *Journal of Official Statistics* vol.3, no.4, pp.449—457.

Berlin, M, Mohadjer, L, Waksberg, J, Kolstad, A, Kirsch, I, Rock, D & Yamamoto, K 1992, 'An Experiment in Monetary Incentives', *1992 Proceedings of the Survey Research Methods Section of the American Statistical Association*, 292—298.

Brownell, MD, Roos, NP, MacWilliam, L, Leclair, L, Ekuma, O & Fransoo, R 2010, 'Academic and social outcomes for high-risk youths in Manitoba', *Canadian Journal of Education*, vol.33, pp.804—36.

Bullen, P 2004, 'Social capital questions for the Longitudinal Surveys of Australian Youth 2004, unpublished.

Burkham, D & Lee, V 1998, 'Effects of monotone and non-monotone attrition on parameter estimates in regression models with educational data: Demographic effect on achievement, aspirations and attitudes', *Journal of Human Resources*, vol.33, no.2, pp.555–74.

Bureau of Labour Statistics 2013, 'What information is available from the National Longitudinal Surveys (NLS)', viewed October 2013, <http://www.stats.bls.gov/nls/nlsfaqs.htm#anch11>.

Bynner, J, Elias, P, McKnight, A, Pan, H & Pierre, G 2002, *Young people's changing routes to independence*, Joseph Rowntree Foundation, viewed October 2013, <http://www.jrf.org.uk/publications/young-peoples-changing-routes-independence>.

Campbell, C, Proctor, H & Sherington, G 2009, *School choice: How parents negotiate the new school market in Australia*, Allen and Unwin, Sydney.

Chowdry, H, Crawford, C, Dearden, L, Goodman, A & Vignoles, A 2008, *Widening participation in higher education: Analysis using linked administrative data*, Institute for Fiscal Studies, London.

Church, AH 1993, 'Estimating the effects of incentives on mail survey response rates: A meta-analysis', *Public Opinion Quarterly*, vol.57, no.1, pp.62–79.

Citro, CF & Kalton, G 1993, *The future of the survey of income and program participation*, National Academies Press, viewed October 2013, <http://www.nap.edu/catalog/2072.html>.

Cobb-Clark, D 2008, *Leaving home: What economics has to say about the living arrangements of young Australians*, Australian National University Centre for Economic Policy Research discussion paper no.568, viewed October 2013, <http://cbe.anu.edu.au/research/papers/ceprdpapers/DP568.pdf>.

Cochran, WG 1977, *Sampling techniques*, 3rd edn, Wiley series in probability and mathematical statistics — probability, John Wiley and Sons, New York.

Coleman, JS 1988, 'Social capital in the creation of human capital', *American Journal of Sociology*, vol.94 (Supplement), S95–S120.

Collingwood, A, Cheshire, H, Nicolaas, G, D'Souza, J, Ross, A, Hall, J, Armstrong, C, Prosser, A, Green, R, Collins, D, Gray, M & McNaughton, NC 2010, 'A review of the Longitudinal Study of Young People in England (LSYPE): Recommendations for a second cohort', viewed September 2013, <http://dera.ioe.ac.uk/864/1/DFE-RR048.pdf>.

Commonwealth Scientific and Industrial Research Organisation & University of South Australia 2008, *2007 Australian National Children's Nutrition and Physical Activity Survey: Main findings*, Departent of Health and Ageing, Australian Food and Grocery Council, and Department of Agricuture, Fisheries and Forestry, Canberra.

Curtis, D & McMillan J 2008, *School non-completers: Profiles and initial destinations*, research report 54, ACER, Melbourne.

Daraganova, G, Edwards, B & Sipthorp, M 2013, *Using National Assessment Program — Literacy and Numeracy (NAPLAN) data in the Longitudinal Study of Australian Children (LSAC)*, Department of Families, Housing, Community Services and Indigenous Affairs, Canberra.

De Vaus, D (ed.) 2002, *Surveys in social research*, Allen and Unwin, Sydney.

Dezateaux, C (work in progress), *Linkage of health records for children in the Millennium Cohort Study*, viewed September 2013, <http://www.ucl.ac.uk/ich/research-ich/mrc-cech/research/studies/linkage-MSC>.

Dix, KL, Slee, PT, Lawson, MJ, & Keeves, JP 2012, 'Implemenation of quality whole-school mental health promotion and students' academic performance', *Child and Adolescent Mental Health*, vol.17, no.1, pp.45–51.

Dorofeev, A & Grant, P 2006, *Statistics for real-live sample surveys*, Cambridge University Press, London.

Doyle, JK 2005, 'Face-to-face surveys', in *The Encyclopaedia of Statistics in Behavioural Science*, eds BS Everitt & D Howell, Wiley, New York.

Fitzgerald, J, Gottschalk, P & Moffat 1998, 'The impact of attrition in the panel study of income dynamics of intergenerational analysis', *Journal of Human Resources*, vol.33, no.2, pp.300–44.

Gemici, S & Nguyen, N 2013, *Data linkage and statistical matching: options for the Longitudinal Surveys of Australian Youth*, technical paper 77, NCVER, Adelaide.

Gemici, S, Lim, P & Karmel, T 2013, *The impact of schools on young people's transition to university*, research report 61, NCVER, Adelaide.

Goleman, D 1995, *Emotional intelligence*, Bantam Books, New York.

Goodman, R 1997, 'The Strengths and Difficulties Questionnaire: A research note' *Journal of Child Psychology and Psychiatry*, vol.38, no.5, pp.581–6.

Gregory, T 2012, 'Early development index (EDI) at age 5 predicts reading and numeracy skills three to seven years later', paper presented at the 2012 International Data Linkage Conference, Perth, viewed December 2012, <http://datalinkage2012.com.au/program/presentations/gregory.pdf>.

Halliday Wynes, S & Nguyen, N 2014, *The impact of financial stress on young people in tertiary study*, NCVER, Adelaide.

Hansen, K, & Joshi, H (ed.) 2007, *Millennium Cohort Study Second Survey: A user's guide to initial findings*, Centre for Longitudinal Studies, London.

Hartley, R 1993, 'Young Adults Living at Home', *Family Matters*, vol.36, pp.35–37.

Harris, KM, Halpern, CT, Whitsel, E, Hussey, J, Tabor, J, Entzel, P, & Udry, JR 2009, *The National Longitudinal Study of Adolescent Health: Research Design*, viewed November 2013, <http://www.cpc.unc.edu/projects/addhealth/design>.

Hockley, C, Quigley, MA, Hughes, G, Calderwood, L, Joshi, H, Davidson, LL 2008, 'Linking Millennium Cohort data to birth registration and hospital episode records', *Paediatric and Perinatal Epidemiology*, vol.22, no.1, pp.99–109.

Howieson, C, Croxford, L & Howart, N 2008, *Meeting the needs for longitudinal data on youth transitions in Scotland: An options appraisal*, Scottish Government, Edinburgh.

Hoyt, LT, Chase-Lansdale, PL, McDade, TW, Adam, EK 2012, 'Positive youth, healthy adults: Does positive well-being in adolescence predict better perceived health and fewer risky health behaviours in young adulthood?', *Journal of Adolescent Health*, vol.50, no.1, pp.66–73.

James, TL 1997, 'Results of the Wave 1 Incentive Experiment in the 1996 Survey of Income and Program Participation', *1997 Proceedings of Survey Research Methods Section of the American Statistical Association*, pp.834–839.

Johnson, S & Antill, M 2011, *Impact evaluation of the millennium cohort study*, Consulting Inplace, London.

Jones, AM, Koolman, X & Rice, N 2006, 'Health-related non-response in the British Household Panel Survey and European Community Household Panel: using inverse-probability-weighted estimators in non-linear models', *Journal of the Royal Statistical Society*, Series A (Statistics in Society), vol.169, pp.543–69.

Jutte, DP, Brownell, M, Roos, NP, Schippers, C, Boyce, WT & Syme, SL 2010, 'Rethinking what is important: biologic versus social predictors of childhood health and educational outcomes', *Epidemiology*, vol.21, no.3, pp.314—23.

Jutte, DP, Roos, LL & Brownell, MD 2011, 'Administrative record linkage as a tool for public health research', *Annual Review of Public Health*, vol.32, pp.91—108.

Kalton, G 2009, 'Methods for oversampling rare subpopulations in social surveys', *Survey Methodology*, vol.35, no.2, pp.124—41.

Kaminska, O & Lynn, P 2013, 'Tailoring mode of data collection in longitudinal studies', paper presented at the HILDA Survey Research Conference, 2013, Melbourne, viewed November 2013, <http://www.melbourneinstitute.com/downloads/conferences/HILDA_2013/HILDA_2013_papers/Kaminska,%20Olena_final%20paper.pdf>.

Karmel, T 2013, *Are we there yet? Overview of the Longitudinal Surveys of Australian Youth*, NCVER, Adelaide.

Kho, ME, Duffett, M, Willison, DJ, Cook, DJ & Brouwers, MC 2009, 'Written informed consent and selection bias in observational studies using medical records: systematic review', *British Medical Journal*, doi:10.1136/bmj.b866.

Knies, G, Burton, J & Sala, E 2012, 'Consenting to health record linkage: evidence from a multi-purpose longitudinal survey of a general population', *BMC Health Services Research*, vol.12, viewed November 2012, <www.biomedcentral.com/content/pdf/1472-6963-12-52.pdf>.

Lawlor, DA, Sterne, JAC, Tynelius, P, Smith, GD & Rasmussen, F 2006, 'Association of childhood socioeconomic position with cause-specific mortality in a prospective record linkage study of 1,839,384 individuals', *American Journal of Epidemiology*, vol.164, no.9, pp.907—15.

Levitt, JM, Saka, N, Romanelli, LH, & Hoagwood, K 2007, 'Early identification of mental health problems in schools: the status of instrumentation', *Journal of School Psychology*, vol.45, no.2, pp.163–91.

Li, X, Sundquist, J & Sundquist, K 2010, 'Parental occupation and risk of small for-gestational-age births: A nationwide epidemiological study in Sweden', *Human Reproduction*, vol.25, no.4, pp.1044—50.

Lim, P, Gemici, S, & Karmel, T 2013, *The impact of school academic quality on low socioeconomic status students*, NCVER, Adelaide.

Logie, H, Hogan, R & Puet, A 2004, *Longitudinal studies of ageing: Implications for future studies*, Australian Institute of Health and Welfare, Canberra.

Longitudinal Study of Young People in England (LSYPE), 2011, viewed 23 September 2013, <http://discover.ukdataservice.ac.uk/series/?sn=2000030>.

Lynch, J 2012, *South Australian early childhood development demonstration project*, viewed 6 November 2012, <http://www.santdatalink.org.au/Demonstration_Projects>.

Mack, S, Huggins, V, Keathley, D & Sundukchi, M 1998, 'Do monetary incentives improve response rates in the Survey on Income and Program Participation?' *Proceedings of Survey Research Methods Section of the American Statistical Association*, pp.529–534.

Marjoribanks, K 2002 *Family and school capital: Towards a context theory of students' school outcomes*, Kluwer, Dordrecht.

——2004, 'Families, schools, individual characteristics, and young adults' outcomes: Social and cultural group differences', *International Journal of Educational Research,* vol.41, pp.10–23.

——2005, 'Family background, adolescents' educational aspirations, and young Australian adults' educational attainment', *International Education Journal,* vol.6, no.1, pp.104–12.

McCulloch, A 2001, 'Psychiatric morbidity and non-ignorable attrition from the British Household Panel Survey', paper presented at the British Household Panel Survey Research Conference, 2001.

National Health and Medical Research Council, Australian Research Council & Australian Vice-Chancellors' Committee 2009, *National statement on ethical conduct in human research*, rev. edn, Canberra.

National Statistical Service 2010, 'High level principles for data integration involving Commonwealth data for statistical and research purposes', viewed November 2012, <http://www.nss.gov.au/nss/home.nsf/NSS/7AFDD165E21F34FDCA2577E400195826?opendocument>.

——2012, Rights, responsibilities and roles of integrating authorities, viewed November 2012, <http://www.nss.gov.au/nss/home.nsf/pages/Data+Integration+Roles+and+responsibilities+of+integrating+authorities?OpenDocument>.

NCVER (National Centre for Vocational Education Research) 2011, *Longitudinal Surveys of Australian Youth 1998 cohort: User guide*, NCVER, Adelaide.

Nguyen N, Cully, M, Anlezark, A & Dockery, AM 2010, *A stocktake of the Longitudinal Surveys of Australian Youth*, technical paper 59, NCVER, Adelaide.

North Carolina Center for Public Health Preparedness, 'Two-stage cluster sampling: general guidance for use in public health assessments', viewed October 2013, <http://cphp.sph.unc.edu/PHRST5/TwoStageSampling.pdf>.

OECD (Organisation for Economic Development and Cooperation) 1996, *OECD employment outlook 1996,* viewed October 2013, <http://www.keepeek.com/Digital-Asset-Management/oecd/social-issues-migration-health/oecd-employment-outlook-1996_empl_outlook-1996-en#page1>.

——2006, PISA technical report, OECD, Paris.

——2009, *PISA technical manual 2009*, Chapter 4: Sample design, viewed 29 October 2013, <http://www.oecd.org/pisa/pisaproducts/pisa2009/48578536.pdf>.

——2012, *PISA 2009 technical report*, OECD, Paris.

——2013, 'Knowing time: How technology innovations in statistical data collection can make a difference in development', paper presented at the 2013 OECD Global Forum on Development, 2013, Paris, viewed November 2013, <http://www.oecd.org/site/oecdgfd/Session%203.2%20-%20GFD%20Background%20Paper%20-%20DRAFT.pdf>.

Polidano, C, Tabasso, D & Zhang, R forthcoming, *Does scored VET in Schools help or hinder access to higher education in Victoria?,* NCVER, Adelaide.

Purcell, K & Elias, P 2004, *Seven years on: Graduate careers in a changing labour market,* The Higher Education Careers Services Unit, viewed October 2013, <http://www2.warwick.ac.uk/fac/soc/ier/research/completed/7yrs2/seven_years_on.pdf>.

Putnam, RT 1993, 'The prosperous community',*The American Prospect,* vol.4, no.13.

Richardson, AJ, Ampt, ES, Meyburg, AH 1995, *Survey methods for transport planning*, Eucalyptus Press, Melbourne.

Ryan, C 2011, *Year 12 completion and youth transitions,* research report 56, NCVER, Adelaide.

Sala, E, Burton, J & Knies, G 2012, 'Correlates of obtaining informed consent to data linkage: respondent, interview and interviewer characteristics', *Sociological Methods & Research*, vol.41, no.3, pp.414—39.

Salovey, P, Hsee, C, & Mayer, JD 2001, 'Emotional intelligence and the self-regulation of affect', in WG Parrott (ed.), *Emotions in social psychology: essential readings,* US Psychology Press, Philadelphia, pp.185–97.

Schneider, J 1999, *The increasing financial dependency of young people on their parents*, SPRC discussion paper No.96.

Seligman, MEP 1998, *Learned optimism: How to change your mind and your life*, 2nd edn, Pocket Books, New York.

Seligman, MEP, Ernst, RM, Gillham, J, Reivich, K & Inkins, M 2009, 'Positive education: positive psychology and classroom intervetions', *Oxford Review of Education,* vol.35, no.3, pp.293–311.

Singer, E 2002, 'The Use of Incentives to Reduce Nonresponse in Household Surveys', eds Groves, RM, Dillman, DA, Eltinge, JL & Little, RJA, John Wiley & Sons, New York, pp.163—177.

Singh, RP, Petroni, RJ & Allen, TM 1994, 'Oversampling in panel surveys', *The Proceedings from the American Statistical Association, Survey Research Methods Section*, pp.674—679.

Slee, PT, Lawson, MJ, Rusell, A, Askell-Williams, H, Dix, KL, Owens, L, Skrzypiec, G & Spears, B 2009, *KidsMatter evaluation, final report*, Beyond Blue, Melbourne.

Smyth, C, Zappala, G & Considine, G 2002, *School to adult life transitions through work and study: A select review of the literature,* Research and Social Policy Team background paper no.4, The Smith Family, viewed October 2013, <http://www.orfeusresearch.com.au/web_images/Background_Paper_4.pdf>.

Soloff, C, Sanson, A, Wake, M & Harrison, L 2007, 'Enhancing longitudinal studies by linkage to national databases: growing up in Australia, the Longitudinal Study of Australian Children', *International Journal of Social Research Methodology*, vol.10, no.5, pp.349—63.

Spera, C, Wentzel, KR, & Matto, HC 2009, 'Parental aspirations for their children's educational attainment: Relations to ethnicity, parental education, children's academic performance, and parental perceptions of school climate', *Journal of Youth and Adolescence,* vol.38, pp.1140–52.

Stone, W, & Hughes, J 2002, *Social capital: Empirical meaning and measurement validity*, research paper no.27, Australian Institute of Family Studies, Melbourne.

Swedish Centre for Epidemiology 2003, *The Swedish medical birth register: A summary of content and quality*, viewed November 2012, <http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10655/2003-112-3_20031123.pdf>.

Switzerland, Transitions from Education to Employment (TREE), viewed September 2013, <http://tree.unibas.ch/en/the-project/>.

University of Manitoba 2012, *Manitoba population health research data repository list of linked databases*, viewed November 2012, <http://umanitoba.ca/faculties/medicine/units/mchp/protocol/media/database_list_colour.pdf>.

US Bureau of Labor Statistics 2011, 'The NLSY97', viewed September 2013, <http://www.bls.gov/nls/nlsy97.htm>.

Watson, N & Wooden, M 2012, 'The HILSA Survey: A case study in the design and development of a successful household panel study', *Longitudinal and Life Course Studies*, vol.3, no.3, pp.369–381.

Watson, D 2003, 'Sample attrition between waves 1 and 5 in the European Community Household Panel', *European Sociology Review*, vol.19, pp.361–78.

Weisberg, HF 2005, *The total survey error approach: A guide to the new science of survey research*, The University of Chicago Press, Chicago and London.

Weston, R, Stanton, D, Lixia, Q & Soriano, G 2001, 'Australian Families in Transition', *Family Matters*, vol.60, pp.12–23.

Whittington, L & Peters, E 1996, 'Economic incentives for Financial and Residential Independence', *Demography*, vol.33, no.1, pp.82–97.

Appendix A

This appendix provides the draft terms of reference for the LSAY review.

LSAY review terms of reference

Issues to be addressed

The review will:

1. make an objective assessment of the value for money to policy makers, researchers and the wider community of continuing the LSAY program, including a review of actual use by policy makers and others in recent years;
2. identify options and timelines for discontinuance and continuance, including at a reduced funding level, noting that the data collection contract with Wallis Consulting runs to 2015;
3. identify the feasibility, implications and cost of proposed enhancements or changes to LSAY, including:
    a. reducing attrition loss, particularly from the initial PISA group;
    b. reconsidering the cycle for starting new cohorts;
    c. extending the age to which cohorts are followed to 30;
    d. linking to other educational databases such as NAPLAN and potentially to administrative databases such as for employment and social security;
    e. reviewing the survey question set — for example, to collect better information on health and wellbeing, building resilience, and earlier influences and experiences;
    f. introducing a parent questionnaire to collect more comprehensive background information on respondents;
    g. introducing supplementary topical surveys, focus groups or other means of enhancing the usefulness of LSAY to policymakers; and
    h. enhancing LSAY's ability to provide information at regional levels, noting sample size limitations.
4. explore ways to improve LSAY to make it a better and more agile policy tool including in areas such as:
    a. more nuanced evidence around disadvantaged groups, place and region;
    b. what works for young people at risk;
    c. use of up to date technology, data frameworks and communication tools including social media (these have evolved a lot since 1995);
    d. an improved profile and communication strategy; and
    e. options for funding of LSAY, whether enhanced or not, including through an NPP and contribution by other stakeholders;

5. engage strongly with stakeholders on these matters; and
6. conclude with an external assessment of the review and recommendations by an experienced person with a relevant policy background, within a cap of $50,000.

The review will be conducted as an internal Departmental review. A steering committee will be established including a small number of representatives from DEEWR[29] and other agencies with a direct interest. The LSAY Strategic Advisory Committee will act as a reference group. Technical support and analysis will be provided by NCVER under contract on matters including:

- analysis of the technical and cost implications of options arising from the LSAY evaluation and other forms of stakeholder input;
- investigation of options for enhancing cohort retention;
- associated modelling and feasibility work; and
- upgraded specifications for enhanced public data such as longitudinal data cubes and an annual 'quick to market' report on LSAY.

The SIMR committee will be kept informed and its views sought as the review progresses, including full details of the review plan, the contract with NCVER and its cost.

The timeframe is for commencement in May and a final report by the external assessor in November 2013.

[29] Now known as the Department of Education.

Appendix B

This appendix provides information on parental questionnaires administered in the PRLS and PISA.

PIRLS 2006

In the PIRLS 2006 survey, with a focus on the literacy achievement of Year 4 (typically an eight-year-old) students, the parental questionnaire sought information on early literacy practices and on family characteristics such as parent education, employment status, occupation, and relative affluence. These are questions that a child in Year 4 is very unlikely to be able to answer, so in the case of PIRLS a parent questionnaire is essential. In 2006, all 40 countries[30] that participated in PIRLS administered the parent questionnaire.

In those countries, 88% of parents returned the surveys but the return rate varied by country. In England, the return rate was 47%, while in the Russian Federation it was 99%. In 29 countries or provinces, the parent questionnaire return rate was more than 90%.

Table B1 Return rate of PIRLS 2006 parental questionnaire by country

Country (or region or grade level)	Not returned	Returned	Total	Return rate
Austria	209	4858	5067	95.9
Belgium Flemish	135	4344	4479	97.0
Belgium French	440	4112	4552	90.3
Bulgaria	137	3726	3863	96.5
Canada Ontario	455	3533	3988	88.6
Canada Quebec	459	3289	3748	87.8
Canada Alberta	822	3421	4243	80.6
Canada British Columbia	973	3177	4150	76.6
Canada Nova Scotia	416	4020	4436	90.6
Chinese Taipei	118	4471	4589	97.4
Denmark	248	3753	4001	93.8
England	2147	1889	4036	46.8
France	328	4076	4404	92.6
Georgia	111	4291	4402	97.5
Germany	1031	6868	7899	86.9
Hong Kong	126	4586	4712	97.3
Hungary	358	3710	4068	91.2
Iceland	885	2788	3673	75.9
Indonesia	67	4707	4774	98.6
Iran	71	5340	5411	98.7
Israel	1462	2446	3908	62.6
Italy	135	3446	3581	96.2
Kuwait	1021	2937	3958	74.2
Latvia	230	3932	4162	94.5

[30] In several of the 40 countries that participated in PIRLS, data from five Canadian provinces and the Flemish and French speaking regions of Belgium were reported separately. In addition to their Grade 4 samples, Iceland and Norway surveyed Grade 5 students. No data are available for the United States. Thus, data are reported for 46 countries, regions or grade levels.

Country (or region or grade level)	Not returned	Returned	Total	Return rate
Lithuania	104	4597	4701	97.8
Luxembourg	366	4735	5101	92.8
Macedonia	137	3865	4002	96.6
Moldova	136	3900	4036	96.6
Morocco	86	3163	3249	97.4
Netherlands	1300	2856	4156	68.7
New Zealand	2270	3986	6256	63.7
Norway	297	3540	3837	92.3
Poland	108	4746	4854	97.8
Qatar	1869	4811	6680	72.0
Romania	86	4187	4273	98.0
Russian Federation	32	4688	4720	99.3
Scotland	1820	1955	3775	51.8
Singapore	117	6273	6390	98.2
Slovakia	152	5228	5380	97.2
Slovenia	261	5076	5337	95.1
South Africa	1426	13231	14657	90.3
Spain	1533	2561	4094	62.6
Sweden	283	4111	4394	93.6
Trinidad and Tobago	422	3529	3951	89.3
Iceland (5th grade)	442	937	1379	67.9
Norway (5th grade)	194	1614	1808	89.3
United States	na	na	5190	na
Total	**25041**	**173995**	**199036**	**87.4**

PISA 2009

From 2006, the PISA survey has included an optional parent questionnaire. In 2009, the main focus of PISA was reading literacy. In that year, the parent questionnaire was administered in 15 of the 65 countries that participated in PISA. The questionnaire sought information on a variety of topics, including the child's early education experience, parents' and child's home literacy practices, parental education, household income, and parents' perceptions of their child's school. These are topics about which children may not have accurate knowledge.

Response rates

An analysis of non-response to the parent questionnaire by country (see Table B2) may provide information about the likely response rate if a similar survey were administered in Australia. Several observations are made of the result presented in the table. Non-response to the identity of the respondent (mother, father or other) is taken as an indicator of the overall non-return rate of the questionnaire. The overall non-return rate across the 15 countries is good, at only 15.3%. However, the non-return rate varies quite markedly by country, from a low of 1.5% (Korea) to a high of 41.1% (Denmark).

Non-response to the mother's education question is indicative of likely responses to questions that might not be answered accurately by 15-year-old students. Non-response to mother's education is very similar to the overall non-return rate in most countries, suggesting that if the questionnaire was returned, this question was answered. Hong Kong is an exception, with a non-return rate of 3.1% but a non-response rate of 17.5% to this question.

Non-response to the household income question is indicative of the likely non-response to sensitive questions. Over all countries, only 1.5% of parents who returned questionnaires declined to answer this question, but there are notable exceptions, with more than 10% of parents from Germany, Hungary, Lithuania and Poland retuning questionnaires without responding to this item. This question was not asked in Italy, perhaps because a low response rate was anticipated.

Table B2 Non-response rate to selected items in the PISA 2009 parent questionnaire by participating country

Country	Parental respondent Missing observations		Mothers education Missing observations		Household income Missing observations		Students in sample
	No.	%	No.	%	No.	%	
Chile	581	10.2	606	10.7	584	10.3	5669
Germany	1872	37.6	1851	37.2	2673	53.7	4979
Denmark	2432	41.1	2470	41.7	2612	44.1	5924
Hong Kong-China	148	3.1	846	17.5	288	6.0	4837
Croatia	517	10.4	570	11.4	813	16.3	4994
Hungary	196	4.3	266	5.8	910	19.8	4605
Italy	3649	11.8	4274	13.8	n/a	n/a	30905
Korea	75	1.5	136	2.7	120	2.4	4989
Lithuania	71	1.6	98	2.2	653	14.4	4528
Macao-China	173	2.9	140	2.4	304	5.1	5952
New Zealand	1194	25.7	1203	25.9	1527	32.9	4643
Panama	693	17.5	771	19.4	875	22.0	3969
Poland	100	2.0	185	3.8	821	16.7	4917
Portugal	1446	23.0	1517	24.1	1998	31.7	6298
Qatar	3129	34.5	3100	34.1	3676	40.5	9078
Total	**16276**	**15.3**	**18033**	**17.0**	**17854**	**16.8**	**106287**

Note n/a = Question not asked.

Response accuracy

While it is expected that 15-year-old students should be able to answer many questions about their families accurately, it is possible that these students may not have accurate knowledge of some aspects of their family's circumstances. Questions about parental education are asked in many surveys and this is shown to be an important predictor of successful youth transitions; for example, completing secondary schooling, participating in tertiary education and training, and finding employment. Because of its importance, it is useful to be sure of the accuracy of student reports of their parents' levels of education.

In Table B3, the self-reported educational attainment of mothers, taken from the 2009 PISA parent questionnaire, is tabulated against the mothers' level of education reported by 15-year-old students in their responses to the student questionnaire. This is done only for the 15 countries that administered the parent questionnaire. In this table, some categories of education had to be combined because of differences in the ways in which the educational attainment questions were asked in the student and parent questionnaires. The combination of categories leads to a more favourable estimate of the accuracy of students' reports of their mothers' levels of education.

In cases where mothers reported less than completion of secondary education, 80% of students concurred with the levels reported by their parents. Similarly, where mothers reported having a bachelor degree or higher, 84% of students agreed. However, where parents reported completion of upper secondary education or a non-tertiary vocational qualification, the levels reported by only 55% of students coincided, with almost one-quarter of students underestimating their mothers' levels of education. The accuracy of

students' estimates of mothers' education was even lower if their mothers had a tertiary VET qualification, with 43% of students agreeing and 32% of students underestimating their mothers' attainment.

The inaccuracy of students' reports adds random variance to this variable and therefore weakens it as a predictor. It is quite likely that previous estimates of the influence of parental education have been conservative and that a low level of parental education is a greater risk factor to successful youth transition than has been believed.

Table B3 Level of self-reported mothers' education (parent questionnaire) by the level reported by student respondents (student questionnaire)

	Mothers' self-reported level of education (parent questionnaire)						
	Less than senior secondary	Senior secondary or non-tertiary VET	Tertiary VET	Bachelor degree or higher	N/A or invalid response	Missing	Total
Less than senior secondary	80.8	24.8	10.5	2.5	35.5	61.5	38.6
Senior secondary or non-tertiary VET	10.6	55.7	21.0	6.5	22.5	18.3	28.5
Tertiary VET	2.0	7.7	42.5	5.2	11.5	4.1	8.8
Bachelor degree or higher	2.4	9.1	22.8	84.2	20.8	8.3	20.0
N/A or invalid response	0.4	0.2	0.4	0.2	3.1	0.5	0.7
Missing	3.7	2.5	2.7	1.3	6.6	7.3	3.4
Total (%)	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Total (N)	**28633**	**37031**	**7870**	**14280**	**14802**	**3671**	**106287**

National Centre for Vocational Education Research

Level 11, 33 King William Street, Adelaide, SA 5000

PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

Phone +61 8 8230 8400 Fax +61 8 8212 3436

Email ncver@ncver.edu.au Web <http://www.ncver.edu.au> <http://www.lsay.edu.au>

Follow us: <http://twitter.com/ncver> <http://www.linkedin.com/company/ncver>