2013 Public Lending Right for New Zealand Authors Survey

Prepared by Statistics New Zealand
for the National Library of New Zeanald

Charlie Dohrman & Asheel Ramanlal
Statistical Methods
28 September 2012

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

New Zealand Government

Contents

1 Objective 3
2 Background to the survey 3
3 Changes from previous survey 3
4 Technical objectives 4
4.1 Type of surveys 4
4.2 Accuracy requirements 4
5 The frame to be used 4
5.1 Target and survey population 4
5.2 Stratification variables 4
5.3 Details of updating 4
6 The sample design details 5
6.1 Option 1: rotating panel design 5
6.2 Option 2: alternating two-year cycle design 6
6.3 Option 3: one-year cycle design 7
6.4 Allocation to strata 7
6.4.1 For public libraries: 7
6.4.2 For tertiary libraries: 8
6.5 Collection and selection procedures used 8
6.6 Estimation procedure 8
6.7 Editing & imputation 9
6.8 Sources of error 9
6.9 Dissemination of outputs 9
7 Appendices 10
7.1 Appendix A: rotating panel design 10
7.2 Appendix B: alternating two-year cycle design 11
7.3 Appendix C: one-year cycle design 12

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

1 Objective

The purpose of this document is to outline the method used to redesign the Public Lending Right Survey. The National Library will be selecting the sample and collecting the information from the survey with the guidance of the details below, allowing them to be self-sufficient if the population of libraries do not dramatically change.

2 Background to the survey

The Public Lending Right for New Zealand Authors scheme compensates New Zealand authors for the use of their books in New Zealand libraries. The scheme provides annual payments to New Zealand authors. This is to compensate the loss of royalty income incurred as a result of their books being made available through New Zealand libraries.

The Public Lending Right survey is run by the National Library of New Zealand and is designed to distribute this compensation amongst authors who have entered their books on the register. The amount each author receives depends on the total number of copies of registered titles. Authors who have an estimated total of 50 copies or more of their titles in New Zealand libraries receive a percentage of the fund. Authors need to register each year so that their titles become active for that year. If they do not register, their titles will not be included in any survey or payment that year. If they register and then die before payment is made, their estate would receive payment in that year and would not be eligible to register in future years (as they are deceased).

3 Changes from previous survey

The Public Lending Right for New Zealand Authors Regulations 2009 limited the design of the survey in 2010.

Regulation 9 'How New Zealand libraries are to be surveyed or sampled', comprehensively states that:
"The chief executive must do a triennial survey of the National Library and 39 other New Zealand libraries to find out how many copies they hold of books by New Zealand authors who entered their names in the register in the year of the triennial survey. In the first year after the triennial survey, the chief executive must do an updating survey of the National Library and the 39 other New Zealand libraries. In the second year after the triennial survey, the chief executive must do an updating survey of the National Library and the 39 other New Zealand libraries."

Section 8 states that:
"the 39 New Zealand libraries surveyed must be those named by the Government Statistician and must be the same 39 libraries for the 3 surveys."

In 2012, an Amendment to the regulations was put through changing how New Zealand libraries are to be surveyed or sampled. Regulation 9 now reads:
"The surveys may be of (a) samples of New Zealand libraries; or (b) all New Zealand libraries. The chief executive must do the surveys in accordance with the advice of the Statistician on (a) how often to do the surveys; and (b) the design of the surveys."

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

This amendment gives Statistics New Zealand (the statistician) a lot more freedom in the design of the survey. This also shifts the responsibility of the survey design entirely up to Statistics New Zealand. To limit Statistics New Zealand's involvement, the design of the survey has two fundamental principles:

1. **Robust:** To cater for the dynamic population;
2. **Unbiased:** To satisfy the needs of a key stakeholder, The New Zealand Society of Authors (PEN NZ Inc.).

4 Technical objectives

4.1 Type of surveys

There are two options on the type of survey used each time it's run:

1. The complete survey, which is a list of all New Zealand titles in the register.
2. The update survey, which is a list of New Zealand titles submitted to the register between 1 January and 30 April in that year. This includes new titles, or titles where an author adequately justifies including a title in a survey again.

4.2 Accuracy requirements

The accuracy has not been stated by the National Library, who is responsible for running the survey. The survey design will be based on previous sample errors, as they were considered acceptable. Sample errors for this survey are not generally released to the public.

The New Zealand Society of Authors (PEN NZ Inc.) approved the changes to the Public Lending Right for New Zealand Authors Regulation 2008 only if the Department could guarantee the reliability of the survey (a guarantee of 90% accuracy). The survey design assumes that this condition implies that relative sampling errors (RSE) of 10% are acceptable.

5 The frame to be used

5.1 Target and survey population

The target population for this survey is all New Zealand libraries with the following exclusions:

1. School libraries, these are for practical reasons.
2. Private libraries that are not open to the general public as a full list of libraries are unavailable.

5.2 Stratification variables

Information on current public libraries will be known by the National Library, including the total holdings. In 2010 the decision to use 'holdings' (where holdings is the total number of books for each district) as a stratification variable was made. This was considered to correlate the most with the number of New Zealand titles in New Zealand libraries (the variable of interest).

5.3 Details of updating

Information on the holdings for the tertiary libraries, consisting of universities and polytechnics, were found on the following links:

Universities: http://www.universitiesnz.ac.nz/aboutus/sc/CONZUL/Statistics

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

Polytechnics: http://wikieducator.org/File:Polystats_2005-2011.pdf

This information will enable the design variable 'holdings' to be updated for the population each year. This is important if libraries form new consortiums, join existing consortiums or amalgamate. It will help to establish if a consortium or amalgamation is large enough to move into the full-coverage stratum. If the full-coverage stratum becomes too large, the constraint for the cut-off between full-coverage and smaller strata will have to be revised. Statistics New Zealand should be involved in the revision to ensure that sample errors do not become too high with the change.

The tertiary libraries appear more static and are thus divided into two strata: Universities and Polytechnics. If the tertiary population becomes more dynamic, then the boundaries for the cut-off on stratum should also be revised. Statistics New Zealand should be involved in the revision to eliminate the possibility of high sample errors.

6 The sample design details

6.1 Option 1: rotating panel design

Initially the population needs to have the design variable 'holdings' updated.

Divide this population into 4 strata:

1. Public libraries with holdings >200,000;
2. Remaining public libraries;
3. University libraries;
4. Polytechnic libraries.

Both large public libraries and university libraries will be made full coverage.

The remaining two strata will have a rotating panel that is kept in for two years. In the panels' first year, the complete survey is used, and in the second year the updated survey is used. In the third year, when the panel is rotated out of the sample, it is also removed from the survey population for that year only, allowing other libraries to be chosen. Each year a new panel is chosen which will have a size that is half the sample size. *See Appendix A for a graphical representation of this process.*

The justification for only sending the updated survey to the panel in the second year is based on the life-cycle of a book. Smaller stratum libraries are less likely to remove a large number of books from their shelves each year, leaving the results from the prior year a good representation of the number of NZ titles held in the current year. This rotation will also decrease large shifts seen in the change of the yearly count of the number of books per NZ title, which is a large concern of the one of the key shareholders involved, PEN NZ Inc.

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

Stratum ($_h$)	Population size (N_h)	Sample size (n_h)	Weight (w_h)	Overall sample error
Public libraries >200,000	11	11	1	
Small public libraries	36	10	3.6	
University libraries	8	8	1	
Polytechnic libraries	20	6	3.3	
TOTAL	**75**	**35**		**3.3%**

The sample size is based on the total sample error of 3.3%. In 2010 the design had a total sample error of 3.55%.

6.2 Option 2: alternating two-year cycle design

Initially the population needs to have the design variable 'holdings' updated.

Divide this population into 4 strata:

1. Public libraries with holdings >200,000;
2. Remaining public libraries;
3. University libraries;
4. Polytechnic libraries.

Both large public libraries and university libraries will be made full coverage.
The remaining two strata will be a sample. The type of survey will alternative each year between complete and updating surveys. The year the sample is chosen, will be the year the complete survey is used, keeping the same sample in for two years. The second year will be only the updating survey. *See Appendix B for a graphical representation of this process.*

The assumption as to why the sample is only sent the updated survey in the second year is based on the life-cycle of a book. Libraries are less likely to remove a large number of books from their shelves each year, leaving the results from the prior year a good representation of the number of New Zealand titles held in the current year.

Stratum ($_h$)	Population size (N_h)	Sample size (n_h)	Weight (w_h)	Overall sample error
Public libraries >200,000	11	11	1	
Small public libraries	36	9	4	
University libraries	8	8	1	
Polytechnic libraries	20	5	4	
TOTAL	**75**	**33**		**3.51%**

The sample size is based on the total sample error of 3.51%. In 2010 the design had a total sample error of 3.55%.

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

6.3 Option 3: one-year cycle design

Initially the population needs to have the design variable 'holdings' updated.

Divide this population into 4 strata:

1. Public libraries with holdings >200,000;
2. Remaining public libraries;
3. University libraries;
4. Polytechnic libraries.

Both large public libraries and university libraries will be made full coverage.
The remaining two strata will be a sample. The type of survey each year will be the complete survey only. *See Appendix C for a graphical representation of this process.*

Stratum ($_h$)	Population size (N_h)	Sample size (n_h)	Weight (w_h)	Overall sample error
Public libraries >200,000	11	11		
Small public libraries	36	5	7.2	
University libraries	8	8	1	
Polytechnic libraries	20	3	6.67	
TOTAL	**75**	**27**		**5.03%**

The sample size is based on the total sample error of 5.03%. In 2010 the design had a total sample error of 3.55%.

6.4 Allocation to strata

6.4.1 For public libraries:

A program called 'Strat' was used to find the cut-off boundary between large and small public libraries. Strat was run with a range of sample errors, continually giving a size of 11 for large public libraries. The Strat results, as well as the 'cumulative of the square root of the frequency method' results were used to confirm boundaries (this method was used in 2010 to establish the stratum boundaries). Both methods gave the same results indicating that the boundary set a reasonable cut-off for public libraries.

Previously, the sample selection method for public libraries was ordered by geographical location and then a systematic sample was taken. Since then, libraries have amalgamated or formed consortiums decreasing the remaining libraries in the small stratum (the consortiums are a group of libraries from a similar region, disrupting the spread of libraries throughout New Zealand). The proportion of south island libraries to north island libraries has also decreased, and to create a sample design that is independent of these shifts, a randomised sample based on the proportion of remaining south island libraries is used.

In 2012, there were 10 south island public libraries in the small strata, a proportion of 10/36 approximate to 3/10. So instead of ordering libraries, they were split into two groups - South Island and

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

North Island, assigned a random number, and 7 North Island public libraries with the lowest random number were selected as well as 3 south island public libraries with the lowest random number.

6.4.2 For tertiary libraries:

All university libraries were bigger (by number of holdings) than all polytechnic libraries, so it was decided to make university libraries full-coverage and sample the polytechnic libraries.

The sampling method for the polytechnic libraries has not changed. Polytechnic libraries will be assigned a random number and 6 will be chosen that have the lowest random number.

A census was not considered for this survey because the cost of enumeration is too high at the present.

6.5 Collection and selection procedures used

The National Library will be operating the collection process. The selection of sample will be clearly outlined, so the National Library will be able to follow and be self-sufficient in the selection process.

Libraries amalgamating or forming consortiums reduces the costs involved in the collection. A search for titles can be completed in one location and obtaining all information from the libraries within the amalgamation or consortium.

6.6 Estimation procedure

The sampling weight for a stratum h, (w_h) is calculated as follows:

$$w_h = \frac{N_h}{n_h}$$

Where
N_h = Population size of stratum
n_h = Sample size of stratum

The total number of each specified title in New Zealand libraries (T_j) is calculated as follows:

$$T_j = \left(\sum_{h=1}^{4} w_h \times t_{hj} \right)$$

Where
j = the specified title
t_{hj} = the total number of a specified title in stratum

Each library is assigned to a stratum and given the stratum weight. The aggregated total will be considered the count of the current year, and the Authors' Fund will split accordingly to the number of qualifying books.

In New Zealand payments to authors are determined by the book rate, which is calculated by dividing the available funds among the number of qualifying books:

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

$$Book\ rate = \frac{Available\ funds}{Total\ qualifying\ books}$$

6.7 Editing & imputation

Given the method of collection, there is no non-response for this survey. Each library surveyed will complete the survey. There is no need for any editing & imputation processes applied to the data. Hence there is, if any, very little non-sampling error.

6.8 Sources of error

The estimates in this report are based on a sample of libraries. Because the estimates are based on a sample of libraries, all estimates have a sampling error associated with them. The variability of a survey estimate, due to the random nature of the sample selection process, is measured by its sampling error. Sampling errors vary from estimate to estimate, and with population breakdown and population size.

The sampling errors have been calculated based on previous data supplied on the 2010 triennial survey.

Additional notation:
$i = 1, \dots, n_h$ where n_h is the number of libraries in stratum h
$h = 1, \dots, L$ where L is the total number of stratums in the sample
$W_h = \frac{N_h}{N}$ which is the stratum weight
$V(\bar{y}_n) = \frac{s_h^2}{n_h}\frac{N_h - n_h}{N_h}$ which is the stratum variance
$E[\bar{y}_h] = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$ which is the stratum mean

The relative sampling error (RSE) of a stratum is calculated by

$$RSE[\bar{y}_h] = \frac{SE[\bar{y}_h]}{E[\bar{y}_h]} = \frac{\sqrt{V[\bar{y}_n]}}{E[\bar{y}_h]}$$

To find the overall RSE the calculation used follows;

$$RSE[\bar{Y}_{st}] = \frac{SE[\bar{Y}_{st}]}{E[\bar{Y}_{st}]} = \frac{\sqrt{\sum_{h=1}^{L} W_h^2\, V(\bar{y}_h)}}{\sum_{h=1}^{L} W_h\ (\bar{Y}_h)}$$

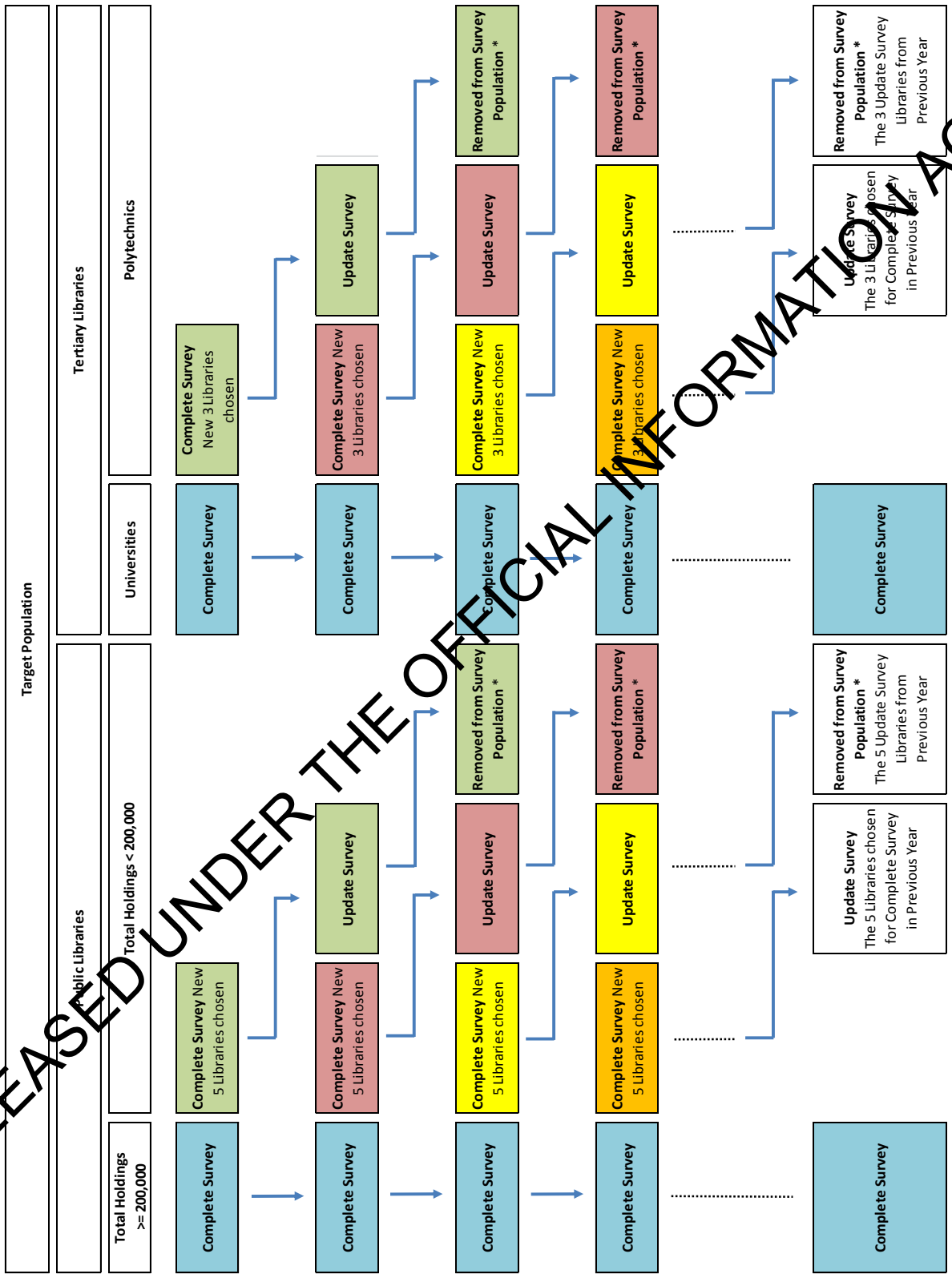
6.9 Dissemination of outputs

Decisions have to be made by the surveying agency (the National Library) on the commercial sensitivity of the data collected. As the survey outputs monetary amounts, unit record data should be adequately confidentialised in order to protect authors.

RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

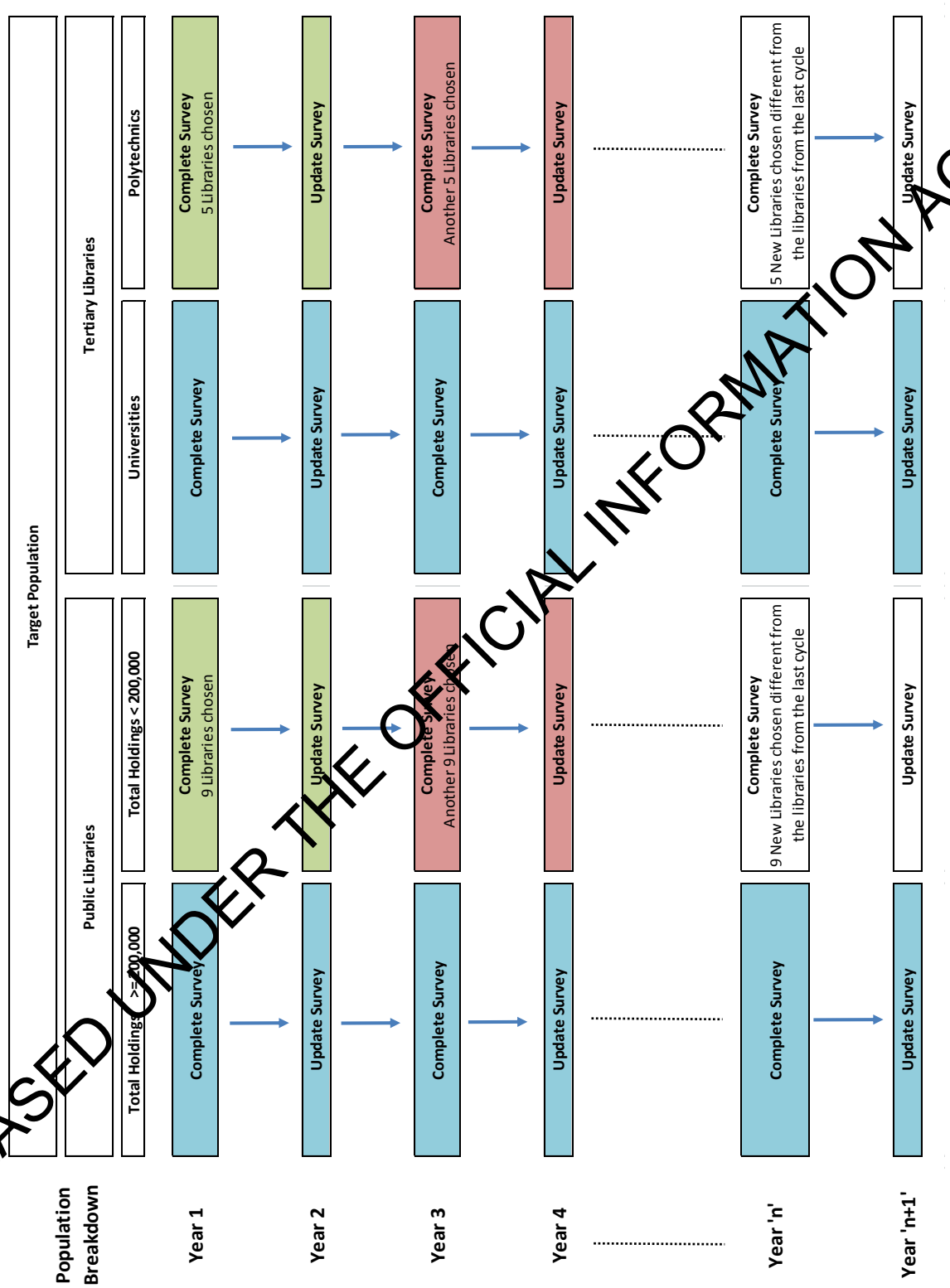
7 Appendices

7.1 Appendix A: rotating panel design



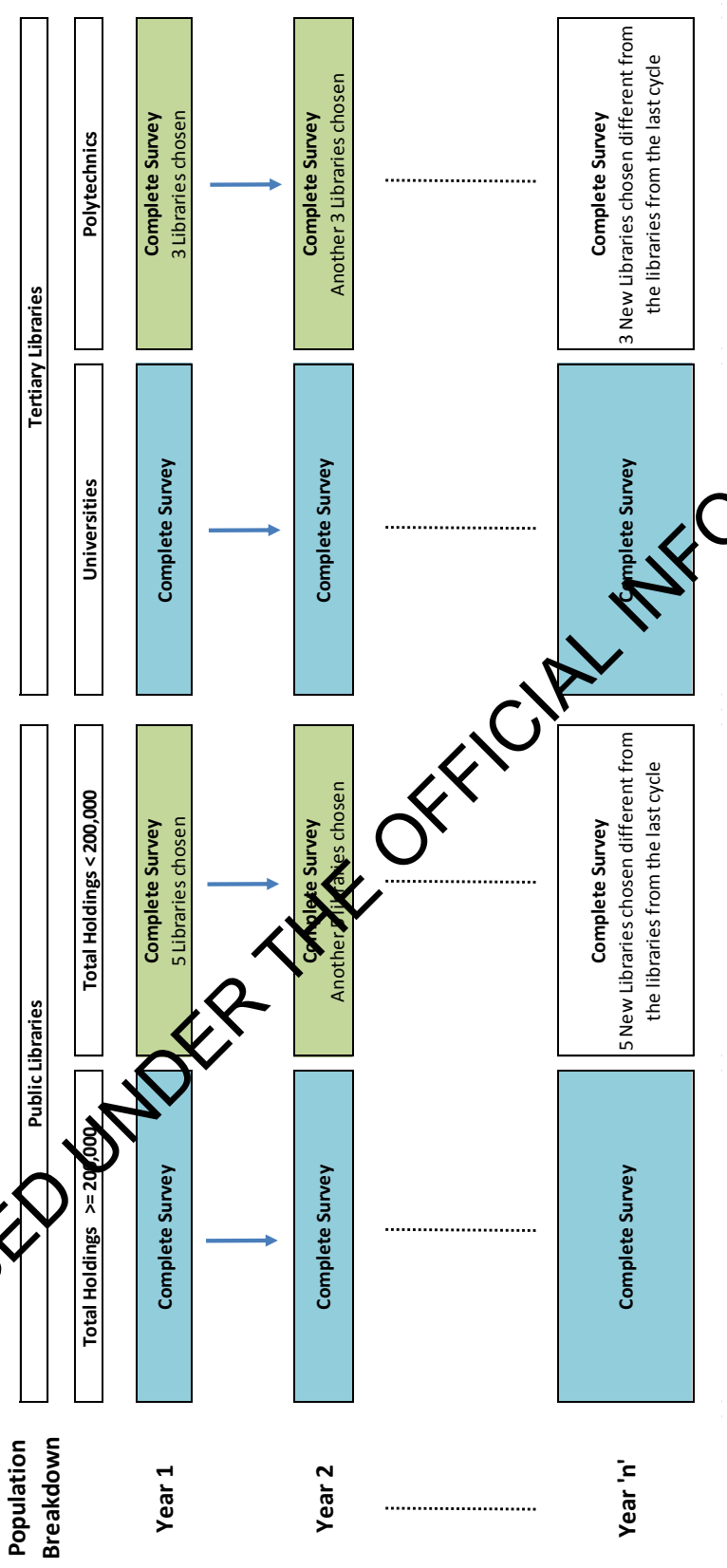
RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

7.2 Appendix B: alternating two-year cycle design



RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982

7.3 Appendix C: one-year cycle design



RELEASED UNDER THE OFFICIAL INFORMATION ACT 1982