



Statistics
Canada

Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca



Project for the Regional Advancement of Statistics in the Caribbean - PRASC

Funded by the
Government
of Canada

Canada



Statistics
Canada

Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca



Project for the Regional Advancement of Statistics in the Caribbean - PRASC

Component: Business Survey Infrastructure

Funded by the
Government
of Canada

Canada



Methodology of Collection Management

Mike Hidioglou and Wesley Yung
Business Survey Methods Division

January 21, 2016
Port of Spain, Trinidad and Tobago



Outline

- Introduction
- Failed Edit Follow-up
- Non-response
- Non-response Follow-up
- Non-response Follow-up Strategies
- Summary



Introduction

- Data collection is recognized as one of the most expensive steps of the survey process
- It is also one of the most important as quality data is needed to produce quality estimates
- Much research has been done in the past decade on how to make the collection process more cost effective
 - For example, responsive design, adaptive designs, electronic questionnaires, etc.



Introduction

- This section will concentrate only on the follow-up process
- There are two types of follow-up typically performed in business surveys
 - Failed edit follow-up (FEFU)
 - Non-response follow-up (NRFU)
- We will present some best practices to make the follow-up process as efficient as possible, keeping quality in mind



Failed Edit Follow-up

- When reported data fail edits (data verification), respondents are re-contacted to confirm or correct the data
- Failed edit follow-up is all about improving the quality of the data
- Re-contacting the respondent is very expensive
 - Human resources required to confirm or correct the data with the respondent
 - Increases the burden on respondents
 - Impacts the good will of the respondent

Failed Edit Follow-up

- Can we reduce the number of failed edits?
 - Computer Assisted Interviewing (CAI)
 - Correct the data while the interviewer is still with the respondent
 - CAI is very expensive, so this collection method is becoming less common
 - Electronic questionnaire (EQ)
 - Put the edits in the EQ and have the respondent fix/confirm the data before continuing
 - Need to be careful to balance quality of the data with the burden placed on respondents
 - Agencies tend to minimize the number of edits in EQ and restrict them to only the most important ones



Failed Edit Follow-up

- Do all failed edits need to be followed up?
 - Probably not
 - Some variables are more important than others
 - Some units are more important than others
- How to decide whether or not to perform follow-up?
 - Can the variable of interest be derived from other information? For example, sum of parts

Failed Edit Follow-up

- Can the variable of interest be imputed using auxiliary data?
 - Historical data
 - Administrative sources such as tax, BR, etc.
 - Alternate sources such as the Internet, other surveys, etc.
- Is the unit so important that the value has to be *exact*?
 - Can we impute a value from another unit?
- Decision to follow-up is a trade-off between cost and quality
- In reality, some units have to be followed-up but certainly not all



Non-response

- Non-response is when some or all data are missing
- Non-response will always occur, but collection management can help to reduce it and its impact on the quality of survey estimates
- There are two types of non-response
 - Partial non-response
 - Unit (or total) non-response

Non-response: Partial

- Some data but not all are collected
- Usually handled through imputation
- There are many ways to impute but this is not in-scope for the workshop
 - Examples include mean, ratio, regression and donor imputation
- The key to quality imputation is availability and use of auxiliary data
- Non-response follow-up usually doesn't target partial non-response

Non-response: Unit

- No data is obtained from a unit
- Impacts the quality of survey results
 - Smaller effective sample size, which leads to a higher variance
 - Potential for non-response bias
- Non-response bias occurs when units that respond no longer represent those which do not
 - For example, suppose businesses with more than 100 employees refuse to respond. Responding units (less than 100 employees) can't be used to represent the non-respondents

Non-Response Follow-Up

- Non-response follow-up is used to reduce the amount of unit non-response
- This in turn, reduces the potential for non-response bias
 - Care must be taken in how it is done, as incorrect non-response follow-up can increase the potential for bias
- While failed edit follow-up targets the quality of the data, non-response follow-up targets the overall quality of the survey results

Non-Response Follow-Up

- Non-response follow-up should be done to ensure that the resulting respondents still represent the non-respondents (as much as possible)
 - The follow-up procedure shouldn't introduce any additional potential for non-response bias
- We present several strategies assuming that the level of follow-up resources is limited

NRFU Strategies

- Follow-up all non-respondents
 - Retains representativity of respondents
 - Follow-up procedure hasn't introduced additional potential for non-response bias
 - May not be that efficient as the level of effort expended on each case may be low
 - Does not allow a targeted follow-up of most important units
 - May lead to misleading estimates
 - Resources are spent on units which may have very little influence on survey estimates

NRFU Strategies

- Follow-up only the important units
 - Appears to be the a logical choice
 - Increases the likelihood of obtaining data from the big businesses which should lead to more accurate estimates
 - Resources are spent where they would make the biggest impact
 - However, sample of respondents is no longer representative
 - Follow-up procedure has increased the potential for non-response bias
 - The impact of this can be offset by imputing for all unit non-respondents
 - Quality of estimates then depend on quality of imputation

NRFU Strategies

- Follow-up important units within each stratum
 - Sample of respondents is more representative than previous strategy
 - Potential for non-response bias is lower
 - Resources are targeted at units to give biggest 'bang for the buck'
 - However, the potential for non-response bias still exists since the respondents in each stratum are not fully representative
 - Can still be offset by imputing for all non-respondents

NRFU Strategies

- Draw a random sample of non-respondents
 - Target the follow-up on these units only
 - Sample of respondents remains representative
 - Follow-up procedure hasn't introduced additional potential for non-response bias
 - Since number of units followed-up can be controlled, the resources expended on each one can be increased
 - How large should the sample be?
 - No guarantee that important units are selected for follow-up
 - Implications for estimation



NRFU Strategies

- Draw stratified sample of non-respondents
 - Same as previous strategy except
 - Important units can be treated a take-all. That is, all important units are followed-up
 - Can sample at different rates according to stratum. Therefore, small (less important) units can be less likely to be selected than medium or large units

Summary

- Follow-up an important and very resource intensive step of the survey process
- A question of trade-offs between cost (monetary and respondent relations) and quality
 - Can achieve very high quality but at what cost?
- Many strategies exist but need to consider potential for non-response bias and not just increasing response rates or targeting the most important units



You can contact the PRASC team at:

prasc@statcan.gc.ca

or

statcan.prasc-prasc.statcan@canada.ca

Canada