STANDARDS AND GUIDELINES

FOR

REPORTING OF NONRESPONSE RATES:

Definitions, Framework and Detailed Guidelines

TABLE OF CONTENTS

Standards and Guidelines
for Reporting of Nonresponse Rates

PREFACE

This document describes the requirements for reporting nonresponse rates in compliance with the Policy on Informing Users of Data Quality and Methodology and for internal analysis.

The standards and guidelines apply to all sample surveys and censuses based on direct data collection from respondents, and to that portion of any programme that involves direct data collection. All of these data collection vehicles are referred to collectively as *surveys* in the remainder of this document. The standards and guidelines apply to unit nonresponse; item nonresponse is not covered.

For purposes of these standards and guidelines, the following definitions are used.

Unit / Item Nonresponse: For the purposes of these standards and guidelines, units are classified as responding or nonresponding. To be classified as responding, the *item response* must meet some threshold level below which the response would be rejected and considered as a *unit nonresponse*.

Integrated Metadatabase (IMDB): A metadatabase describing all statistical activities, which is maintained by Standards Division.

Reporting Requirements

The requirements for reporting of nonresponse information for the purposes of informing users of data quality are governed by the Policy on Informing Users of Data Quality and Methodology. Users should be informed regularly of the nonresponse rate at the estimation phase, the unit on which the rate is based, and whether or not the rate is weighted, for all occasions of the survey. It is suggested that refusal and no contact rates should also be reported when these may be relevant to users in assessing data quality. In addition to reporting of nonresponse rates to users, the weighted and unweighted response rate at the estimation phase should be recorded on the IMDB.

Program areas are responsible for reporting of nonresponse information. The Methods and Standards Committee is responsible for initiating periodic evaluations of the state of compliance with the standards.

Inquiries relating to the interpretation of these standards and guidelines should be addressed to the Chairperson of the Methods and Standards Committee.

1. INTRODUCTION

Information on survey response and nonresponse can serve multiple purposes and multiple users. The data can be used to inform users of data quality, monitor survey operations, evaluate data collection strategies, provide measures of frame coverage, analyse trends in nonresponse over time and across surveys and develop methods to reduce nonresponse. Different users have differing needs. Survey data users will focus on data quality, data collectors need to know the cost of survey operations and survey divisions need to evaluate both data quality and collection costs.

This paper provides a standard conceptual framework for defining and measuring response and nonresponse. The framework is flexible to meet the varying needs of individual surveys. It allows a choice of units for which weighted or unweighted rates may be calculated.

Standardized methods of calculating these rates are desirable to permit comparison between surveys and analysis of trends in data collection and respondent behaviour over time and across surveys. Collection of nonresponse information will be coordinated with reporting of information for the Integrated Metadatabase (IMDB). In addition, it is recommended that each survey maintain a longitudinal database that tracks nonresponse information over time. Detailed nonresponse tracking will facilitate more powerful analysis allowing, for example, for early detection of trends that may not be evident when looking at surveys individually and at specific time periods.

Guidelines are presented for use of subsets of the standard measures for two different purposes.

The first purpose is the reporting of information on nonresponse to users required by the Policy on Informing Users of Data Quality and Methodology. Nonresponse measures are needed to assist users in assessing the "fitness for use" of survey data. These are basic rates, which are currently being produced for most surveys but not always in a consistent and coherent manner. Guidelines for the presentation of information on nonresponse bias to be reported are also discussed under this Policy.

The second purpose is the monitoring of survey operations, and evaluation of data collection methodology and frame coverage. For this, the guidelines consider more detailed measures and breakdowns of the standard measures. While some of these measures are currently being produced for certain surveys, others are beyond the scope of what is currently possible. In this sense the guidelines are to be viewed as goals to be aimed for. In the years ahead, as data collection operations become increasingly automated through the use of generalized collection and capture software, more reporting of these measures will be possible.

1.1 Scope

The standards apply to censuses and sample surveys which are based on direct data collection from respondents. They do not apply to surveys based only on administrative records. However, for surveys based on direct data collection for some units and administrative records for other units, the standards apply to the portions of the survey based on direct data collection.

The standards consider nonresponse at two phases - the data collection phase and the estimation phase. An appropriate framework and rates are defined for each of these phases.

The standards apply to unit nonresponse. That is, the standards are based on classifying units as responding, nonresponding, etc. To be classified as responding, the item response must meet some threshold level below which the response would be rejected and considered as a unit nonresponse. Units that file only a partial response to a survey may be classified as •responding• if the omitted items can be estimated or imputed for in a reliable fashion. If, however, important aggregates or disaggregations such as geographic or industrial detail are missing it will be necessary to record this fact and to calculate a partial non-response rate.

The standard methods can be used to calculate either unit or item nonresponse rates. The primary purpose is the presentation of unit rates. These can be based on a key control variable chosen by the survey manager. When warranted, response rates for other variables may be measured and published.

Related measures such as response burden, adequacy of survey coverage, length of interview process, and survey cost breakdowns are beyond the scope of these guidelines. Although such measures clearly have a place beside nonresponse measures in the analysis of participation in surveys they need to be dealt with separately.

2. STANDARD FRAMEWORK AND DEFINITIONS

Two frameworks are required, one for the data collection phase, the other for the estimation phase.

The framework for the data collection phase is presented first, followed by the definition of the rates applicable to this phase. The estimation framework and the rates pertaining to the estimation phase are described in terms of their differences from the data collection phase.

Before considering the framework for the data collection and estimation phases, it is necessary to describe first the choice of the unit on which to define the nonresponse rates and then how weighted and unweighted rates are calculated and why these different rates are of interest.

2.1 Choice of Unit

The unit for measurement and reporting of nonresponse rates can be the sampling unit, the collection unit, or the unit on which analysis is based. The choice of reporting unit is complicated by the differing needs of participants in the survey process. Groups, or divisions, focusing primarily on data collection are interested in measures of work effort and associated costs. They are likely, therefore, to focus on collection units and to regard any out-of-scope units as having been resolved. Data users, who are more interested in estimates of the statistical properties of the data, may prefer estimates relating directly to the statistical unit for which data are reported. In complex business enterprises, the collection unit and the statistical unit can differ and this may lead to difficulties in linking response rates at the two levels. Finally, survey managers, who are ultimately responsible for the cost and the content of any statistical program are likely to be interested in non-response data for both collection and statistical units.

Because of the complexity of any business universe, the survey specific collection units that are

defined on the Business Register (BR), are the preferred units for business surveys when calculating response rates at the collection phase. Surveys, which use the Collection Entities generated by the BR, should base their response rate calculations upon the collection units. Similarly, surveys using establishment samples generated by the BR should use these BR sampling units for calculating response rates. Business surveys that use non-BR survey units should attempt to measure their response rates on a basis that produces the most comparability with related surveys using the BR At the estimation phase the statistical unit is recommended but it is recognised that this may not always be feasible because most of the pertinent information is gathered at the collection unit level.

For household surveys, the sampling unit generally ought to be used as the basis for information on nonresponse; that is, the dwelling for surveys based on area sampling or on dwelling lists, and the telephone number for telephone surveys (though for the latter some adjustment will be required to allow for households with more than one telephone). In some cases where both the dwelling and person are sampling units for certain subsets of information on nonresponse, both dwelling and person counts should be used in the calculation of the rates.

2.2 Weighting Method

Nonresponse rates can be either weighted or unweighted, and may be calculated for either the collection unit or the statistical unit. In general these guidelines can be more easily applied to collection units. When collection units and statistical units differ, calculating weighted response rates for statistical units may be considerably more complex than calculating weighted response rates for the collection unit. Indeed, while most calculations in these guidelines can be carried out directly by obtaining data generated as a by-product of the data collection process, it may be necessary to create a subsidiary survey process so that nonresponse rates can be calculated for statistical units.

For business surveys, due to the highly skewed nature of the populations, where a relatively small number of units can account for a large portion of the economic activity, nonresponse ought to be calculated both on a weighted and unweighted basis.

In the unweighted case, all units contribute equally to the rates. The nonresponse rates are based simply on the number of units in the different categories required under the definition.

In the weighted case, there are two distinct ways to calculate weighted nonresponse rates. The first option is to define weights in terms of the sampling weights in order to provide an estimate of the population response rate. Such a rate provides a measure of the proportion of the population units that would not respond if the entire population were enumerated.

Alternatively, a weighted rate that takes into account both the sampling weight and the expected value for a variable or activity of interest can be defined. Such a rate provides a measure of the proportion of the total activity missing due to nonresponse. The expected value of the variable of interest can be based on (i) historical data for the variable of interest available from previous rounds of the survey, (ii) from a recent census or (iii) a proxy for the variable of interest obtained from some other source such as administrative records. In business surveys this type of weighted nonresponse rate is most appropriate For example, it is quite possible that missing a single large firm can be responsible for 20% of the industry sales being unreported. On the other hand, missing one of the

smallest units may have no measurable impact on the industry estimates. For very small units this is often true even when the survey weight is accounted for.

Unweighted rates are of interest for monitoring of the data collection processes of both household and business surveys.

Weighted rates are often more useful when assessing the quality of the data outputs. In the case of business surveys, weighted rates based on a key variable at the estimation phase should be used in reporting information on nonresponse and the expected effect of nonresponse on data provided to users.

Household surveys are not as highly affected by skewed populations as are business surveys and weighted rates based on the sampling weights are preferred.

2.3 Data Collection Phase

2.3.1 Data Collection Phase: Framework

The framework is based on the classification of units included in the survey into a nested hierarchy of 15 categories, displayed in box chart format in Figure (1), as follows:

2.3.2 Data Collection Phase: Definition of Units to Calculate the Rates

(1) Total Units. This category corresponds to all units included in the census or sample survey.

(2) Resolved Units. These are the units whose status has been resolved by the end of the period of survey data gathering, as either belonging or not belonging to the target universe for the survey.

(3) Unresolved Units. For many surveys this category does not apply, since all units are resolved. For other surveys, however, it is either impossible or impractical to resolve all units. For example, in a telephone survey of households there are telephone numbers that ring but do not correspond to working numbers. Without checking the status of each so-called ring-no-answer case with the telephone company, there is no way to determine whether such a number represents a working number. Similarly for a business survey with mail collection, without follow-up of units not returning a questionnaire, it may not be known which units are out-of-scope (e.g., the business no longer exists, or it exists but it is in another industry), versus those which are in scope and should have responded.

(4) In-Scope Units. These are units, which have been determined to belong to the target universe for the survey.

(5) Responding units. This category includes all units, which, at the data collection phase, are deemed to have responded by virtue of having provided usable information. At this phase, survey-specific thresholds for "usable information" may be defined in terms of the level to which incomplete questionnaires have to be filled out before a unit is classified as a respondent.

(6) Refusal Conversions. These are refusals from the current or the previous collection period that are successfully converted to be respondents as a result of special refusal conversion efforts. For example, in some monthly surveys, units refusing in one month are referred in the next month to senior interviewers or other interviewers specializing in refusal conversions. Where such efforts are successful, the unit counts as a refusal conversion for the survey round in which it is converted, and would be so classified for only that round.

(7) Other Responding Units. These are the Responding Units that are not Refusal Conversions (See Figure 1).

(8) Nonresponding units. These are in-scope units that are either nonrespondents or that provide no usable information as defined in (6) above.

(9) Refusals. These are nonresponding units that have been contacted but refuse to participate in the survey.

(10) No Contacts. These are in-scope units that cannot be contacted. For household surveys, these include dwellings whose occupants were temporarily absent and households where no one was at home when interviews were attempted. The occupancy status of such dwellings is determined through observation, or where applicable by speaking to building superintendents. For business surveys, these include telephone respondents who cannot be reached, and mail non-respondents who were not contacted as part of any non-response follow-up.

(11) Residual Nonresponding Units. These include units that did not respond due to special conditions (for example, language problems, or inaccessibility) as well as respondents who provided no usable information. Special conditions also include in-scope units for which interviews were not attempted to avoid unwanted overlap between samples for different surveys, as a measure to prevent undue respondent burden. While these latter units differ from other nonresponse in that interviews are not attempted, it is important that they be considered as nonrespondents in deriving nonresponse adjustment factors.

(12) Out-of-Scope Units. These are units included in the survey, but which have been determined to not belong to the target universe for the survey. Such units occur because of the imperfect nature of the sampling frames used in practice. For example, in household surveys in which dwellings are the sampling unit, vacant dwellings are out-of-scope. Similarly, in telephone surveys of households, telephone numbers not corresponding to working residential numbers are out-of-scope. For a business survey of a particular industry, out-of-scope units include those that have gone out of business, and those businesses belonging to another industry. In addition, for some business surveys, seasonal businesses in the off-season are also out of scope.

(13) Non-existent Units. These are the portions of out-of-scope units that were determined to be non-existent; for example, businesses no longer in existence ("deaths") and dwellings that have been demolished.

(14) Temporarily Out-of-Scope Units. These are units which were out-of-scope at the time of the survey, but which might be in-scope at a later date. Hence, units can be temporarily out-of-scope even for one-time surveys. For recurring surveys, it is necessary to recontact temporarily out-of-scope cases periodically in case their status has changed. Examples include businesses that are inactive due to seasonal factors, seasonal dwellings whose occupants have a usual place of residence elsewhere, and vacant dwellings.

(15) Permanently Out-of-Scope Units. These units are out-of-scope because of improper classification for the subject of the survey or because of changes in the classification since the date of the frame. These cases may be screened during the first stage of response.

2.3.3 Data Collection Phase: Definition of rates

Below, we define a number of rates based on the framework and hierarchy presented in Figure 1. In later sections, the rates that should be used for different purposes stated in the Introduction are considered.

(ii) Resolved Rate

$$= \frac{\text{Resolved Units}}{\text{Total Units}}$$

$$= \frac{\text{Box (2)}}{\text{Box (1)}}$$

(iv) In-scope Rate

$$= \frac{\text{In-scope Units}}{\text{Resolved Units}}$$

$$= \frac{\text{Box (4)}}{\text{Box (2)}}$$

(v) Response Rate for data collection

$$= \frac{\text{Responding units}}{\text{In-scope Units + Unresolved Units}}$$

$$= \frac{\text{Box (5)}}{\text{Box (4) + Box (3)}}$$

(vi) Refusal Conversion Rate

$$= \frac{\text{Refusal Conversions}}{\text{Refusals + Refusal Conversions}}$$

$$= \frac{\text{Box (6)}}{\text{Box (9) + Box (6)}}$$

(viii) Nonresponse Rate for data collection

$$= \frac{\text{Nonresponding units + Unresolved Units}}{\text{In-scope Units + Unresolved Units}}$$

$$= \frac{\text{Box (8) + Box (3)}}{\text{Box (4) + Box (3)}}$$

(ix) Refusal Rate

$$= \frac{\text{Refusals}}{\text{In-scope Units}}$$

$$= \frac{\text{Box (9)}}{\text{Box (4)}}$$

(x) No Contact Rate for data collection

$$= \frac{\text{No Contacts + Unresolved Units}}{\text{In-scope Units + Unresolved Units}}$$

$$= \frac{\text{Box (10) + Box (3)}}{\text{Box (4) + Box (3)}}$$

(xi) Residual Nonresponse Rate

$$= \frac{\text{Residual Nonresponding Units}}{\text{In-scope Units}}$$

$$= \frac{\text{Box (11)}}{\text{Box (4)}}$$

(xii) Out-of-scope Rate

$$= \frac{\text{Out-of-scope Units}}{\text{Resolved Units}}$$

$$= \frac{\text{Box (12)}}{\text{Box (2)}}$$

(xiii) Non-existent Rate

$$= \frac{\text{Non-existent Units}}{\text{Resolved Units}}$$

$$= \frac{\text{Box (13)}}{\text{Box (2)}}$$

(xiv) Temporarily Out-of-Scope Rate

$$= \frac{\text{Temporarily Out-of-Scope Units}}{\text{Resolved Units}}$$

$$= \frac{\text{Box (14)}}{\text{Box (2)}}$$

(xv) Permanently Out-of-scope Rate

$$= \frac{\text{Permanently Out-of-scope Units}}{\text{Resolved Units}}$$

$$= \frac{\text{Box (15)}}{\text{Box (2)}}$$

2.4 Estimation Phase

2.4.1 Estimation Phase: Framework

There are two differences between the framework for the data collection and that for the estimation phase.

The first difference relates to the treatment of the category "Unresolved Units". At the data collection phase, these are effectively treated as a category of No Contact Nonresponse. However, at the estimation phase, there should be an attempt to estimate the proportion of the unresolved units that are in-scope and hence should be treated like nonrespondents as opposed to the proportion of the unresolved units that are out-of-scope. This distinction is particularly important if the survey estimates are adjusted to compensate for nonresponse.

The second difference relates to the treatment of those responding units at the data collection phase that provided no usable data for estimation purposes and hence, should be re-classified as nonresponding units.

2.4.2 Estimation Phase: Definition of Units to Calculate the Rates

Hence, for the estimation phase, we have the following new categories (displayed in Figure 2) over and above those pertaining to the data collection phase:

(3A) Estimated In-scope Units. This category estimates the number of units that are in-scope, from among the unresolved units. The proportion of in-scope units among unresolved units can be estimated by following-up a sample of such cases. For recurring surveys, such follow-up studies need only be conducted periodically. For surveys where such follow-up studies are not possible, the estimates might be based on results for other surveys. For surveys where an estimate cannot be made reliably, an assumption needs to be made about the number of unresolved units that are in-scope versus out-of-scope. When there are unresolved units at the data collection phase, this is an important category as it enters into both the numerator and denominator of the nonresponse rate at the estimation phase.

(3B) Estimated Out-of-scope Units. These are the Unresolved Units that are not Estimated In-Scope Units (See Figure 2).

(5A) Responding Units - Unusable. These are units that were considered respondents at the data collection phase, but for estimation purposes, were discovered to be unusable because the level of completion of the questionnaire was below the survey-specific threshold. The category includes also those units that were dropped because errors noted at processing steps such as edit and imputation and problems in coding could not be resolved in time for estimation. For estimation purposes, these units are treated as though they were unit nonrespondents.

2.4.3 Estimation Phase: Definitions of Rates

Hence, for the estimation phase, we have the following revised definitions of three rates from those pertaining to the data collection phase. These include (v), the Response Rate, (viii), the Nonresponse Rate and (x), the No Contact Rate. These rates are defined as follows:

(v) Response Rate for estimation

$$= \frac{\text{Responding units - Responding Units Unusable}}{\text{In-scope Units + Estimated In-scope Units}}$$

$$= \frac{\text{Box (5) - Box (5A)}}{\text{Box (4) + Box (3A)}}$$

(viii) Nonresponse Rate for estimation

$$= \frac{\text{Nonresponding Units + Responding Unusable + Estimated In-scope}}{\text{In-scope Units + Estimated In-scope Units}}$$

$$= \frac{\text{Box (8) + Box (5A) + Box (3A)}}{\text{Box (4) + Box (3A)}}$$

(x) No Contact Rate for estimation

$$= \frac{\text{No contacts + Estimated In-scope Units}}{\text{In-scope Units + Estimated In-scope Units}}$$

$$= \frac{\text{Box (10) + Box (3A)}}{\text{Box (4) + Box (3A)}}$$

2.5 Nonresponse for Secondary or Longitudinal Surveys

A number of surveys are secondary surveys or supplements to a primary survey. These are surveys whose sample consists of respondents to a primary survey. Data collection for the secondary survey may occur at the same time as that for the primary survey, or at a different time. Secondary or supplementary surveys are regularly conducted in association with the Labour Force Survey.

For secondary surveys, nonresponse information should be based on the cumulative nonresponse to both the primary and secondary survey, and to ensure additivity over the two stages, the unit in which nonresponse is measured should be the same for both the primary and secondary survey.

This is an especially important issue in the case of longitudinal surveys. Response rates within a specific time panel will be higher than the response rate across the time panels, since long time nonrespondents may have been eliminated from the roster. In longitudinal surveys there is a greater potential for inconsistencies and overstated response rates. The cumulative nonresponse rate across all time panels is the appropriate measurement tool.

2.6 Nonresponse Bias

Nonresponse is a potential source of bias in survey estimates. Typically, procedures to adjust estimates to compensate for nonresponse are based on a "missing at random" assumption. However, if nonrespondents differ from respondents with respect to the characteristic of interest, then a nonresponse bias will exist.

There are a number of ways to study nonresponse bias. Three possible approaches are described below:

2.6.1 Nonresponse Follow-up

First, a subsample of nonrespondents is subjected to more intensive follow-up. Because of the expense involved, the follow-up may involve a collection of only a subset of the information regularly collected by the survey, providing a measure of nonresponse bias for only a few key variables. Results from the follow-up study can be used to test the validity of the "missing at random" assumption.

2.6.2 Use of Available Information on Nonrespondents

In certain situations, information may be available on some nonrespondents. For example, in a panel survey, units that respond on some but not all occasions are a source of information on characteristics of nonrespondents. For example, characteristics of respondents to a given round who are nonrespondents to the following round may be compared to characteristics of those who respond on both occasions. In other circumstances, certain information may be observable about nonrespondents, for example, in a household survey based on face-to-face interviewing, interviewers may record characteristics of the dwelling, which might be of use in developing a rough profile of nonrespondents and making a plausible adjustment for nonresponse. Another means of obtaining information on characteristics of nonrespondents is through record check studies, in which nonrespondents, and preferably respondents as well, are linked to data from another source - either

administrative data or data from another survey or census. An example would be micro-linkage of household survey nonrespondents and respondents to the Census of Population.

2.6.3 Weighted vs. Unweighted Nonresponse Rates

For surveys where there is no adjustment to compensate for nonresponse in estimation, the comparison of nonresponse rates weighted by a unit size variable with unweighted rates also gives an indication of nonresponse bias in terms of that size variable. For surveys where nonresponse adjustments are made, the weighted nonresponse rates still indicate the portion of the estimate that is at risk to the potential bias of the nonresponse adjustment procedure.

3. MONITORING OF SURVEY OPERATIONS, AND EVALUATION OF DATA COLLECTION METHODOLOGY AND FRAME COVERAGE

3.1 Survey Operations

In monitoring survey operations, it is important to have nonresponse information available while the survey is in progress, to facilitate management of the data collection operations. At certain pre-specified points during data collection operations, interim rates should be produced for this purpose. The ability to produce such rates will be greatly enhanced with the increased automation of data collection and the associated management control facilities.

Rates required for monitoring of survey operations at the data collection phase are the resolved rate, response rate, nonresponse rate, refusal rate, no contact rate, and residual nonresponse rate. The three rates defined at the estimation phase are also required for monitoring, when such rates are relevant.

The resolved rate is a very important measure of the completeness of data collection operations while they are in progress. This rate needs to be complemented by interim nonresponse and response rates that are indicative of the outcome of completed cases. The interim nonresponse and response rates can provide a useful input in later stages of data collection, along with information on other factors such as cost and time, in taking decisions on the appropriate point to cut off data collection operations.

3.2 Evaluation of Data Collection Methodologies

As data collection methodologies become increasingly automated, data generated by the management control systems will provide a wealth of information that can be used to evaluate and improve data collection strategies.

In this environment, response and nonresponse measures that could be produced are described below, followed by a brief discussion of each:

3.2.1 Percentage of Response by Mode of Collection

These rates are clearly relevant only to surveys with two or more modes of collection, as a measure of how much each collection mode has contributed to the overall response rate. Such rates, together with data on costs per case by collection mode, can be used to assess the cost effectiveness of the allocation of sample to the different modes in terms of contribution to the response rate. Care must be taken in interpreting this measure since the choice of method may be correlated with the difficulty of getting response thus confounding the comparison.

3.2.2 Measures of Effort

Just as interim response and nonresponse rates are important in monitoring survey operations, response and nonresponse rates by number of attempts at resolution are useful in assessing the

strategy used by a survey in following up unresolved cases. Such rates and counts will show the incremental increase in response rates resulting from successive follow-ups. For example, they will give the marginal increase in response rate in going from "x" to "x+1" follow-ups. Together with data on costs, these data can be used to assess the cost effectiveness of the follow-up strategy.

The effects on survey estimates by number of follow-ups can be studied so that the follow-up strategy is made optimal in terms of cost and mean squared error. Studies where additional nonresponse follow-up is done for a limited period of time would also permit profiling of the characteristics of nonrespondents and give some idea of the potential nonresponse bias.

The average number of attempted contacts per unit provides a measure of the amount of effort expended in data collection for the survey.

3.2.3 Outcome Rates

Outcome rates can be defined in terms of the percentage of contact attempts resulting in a particular outcome. For example, the resolution outcome rate is the percentage of attempts resulting in the case being resolved. Refusal outcome rates can be defined either unconditionally as the percentage of contact attempts resulting in refusals, or conditionally as the percentage among attempts where contact has been made but the unit refuses.

Outcome rates can be analysed by day and time of day during the collection period, in order to discover whether there are better or worse times for interviewing. Results of such analysis, subject to operational constraints on when interviewing has to take place, can be used to improve scheduling.

3.2.4 Response Rates by Number of Occasions in the Sample

For recurring surveys where units are retained in the sample for two or more occasions, rates by number of occasions in the sample are of interest. Such rates provide an indication of how respondents react to the response burden imposed by successive rounds of the survey, for example, by showing how refusal rates increase as the number of occasions in the sample increases.

3.3 Measures of Frame Coverage

Units which are included in the frame for a survey but which are found to be out-of-scope provide an indication of the overcoverage of the frame. The implications of the overcoverage differ depending on the type of frame. For household surveys based on area sampling, the overcoverage is by design, as described further below. For business surveys based on a list frame, the overcoverage may be symptomatic of both overcoverage and undercoverage in the frame, although data collection operations for a given survey only provide measures of the overcoverage.

For household surveys, when dwelling lists are being created for selected areas, the procedure is generally to list all structures except those that are clearly unfit for human habitation. All the listed dwellings stand a chance of selection, and, once selected, rigorous criteria are applied to classify

them as in-scope or out-of-scope. This exhaustive approach to listing is adopted to reduce the undercoverage of persons, compared with what would occur if more dwellings were excluded at the listing stage. The proper classification of dwellings as in-scope versus out-of-scope (e.g., occupied with residents temporarily away vs. vacant) is important, because erroneously classifying units as out-of-scope will contribute to the undercoverage for the survey. Due to the importance of this classification, periodic checks on how well it is being done ought to be carried out. In the Labour Force Survey, for example, there is a regular vacancy check program. Also a vacancy check program is conducted as part of the Census of Population.

For business surveys, the out-of-scope rate and its further breakdown are important as measures of frame overcoverage. Feedback on the changes in the status of units is also an important input into frame maintenance. Mechanisms have been established such as the Company Contact System for systematic collection of such input.

3.4 Creation of a Longitudinal Database

To facilitate the analysis of nonresponse patterns it is recommended that survey managers create and maintain a longitudinal database of as many nonresponse counts as possible from Figures 1 and 2. Additionally, these items should be broken down into finer detail, for example by industry or by province. Such a database would allow survey managers to identify trends and/or uncover weaknesses in their frames, etc.

4. REPORTING REQUIREMENTS

4.1 Guidelines for Reporting of Nonresponse to Users

The Policy on Informing Users on Data Quality and Methodology is intended to provide users with the information required to assess the "fitness for use" of survey data. Nonresponse rates and information on nonresponse bias are important measures of the completeness and representativeness of the survey results.

These guidelines detail the information to be reported on nonresponse and nonresponse bias under the Policy. As stipulated in the Policy, the information may be reported in the publications containing the data, or by means of references to where the information can be obtained.

4.1.1 Nonresponse

The nonresponse rate at the estimation phase as defined above should be reported, along with the definition used, and an explanation of any deviations from the standard definition given above. Additional details such as the refusal rate or the no contact rate at the data collection phase could be reported in cases where these rates may be relevant to users in assessing data quality, along with the definitions used.

For sub-annual surveys, sub-annual rates as well as annual average rates (by summing the numerators

and denominators of the sub-annual rates) for the preceding calendar year are to be reported. For other surveys, rates for the current round of the survey are to be reported. If preliminary and revised estimates are produced for a survey, the nonresponse rates should be published for both.

The reporting must include the unit in which nonresponse is measured, an indication of whether nonresponse is weighted or unweighted, and if weighted, a description of the weighting method. The methods used to compensate for nonresponse should be described.

A reference or contact should be cited for obtaining more detailed information on nonresponse.

4.1.2 Nonresponse Bias

Nonresponse is a potential source of bias in survey estimates. For recurring surveys there should be periodic studies of nonresponse bias. Findings from these studies should be included in the information reported to users under the Policy.

Information reported under the Policy should also state whether or not survey estimates are adjusted to compensate for nonresponse, and if the estimates are adjusted, a description of the adjustment procedures.

4.2 Reporting requirements to the Integrated Metadatabase

Divisions are asked to calculate and report the response rate to the IMDB. They are required to report response rates for surveys or censuses at the estimation phase, as defined in the standards, both weighted and unweighted. Household surveys are to weight using the survey weights while business surveys are to weight using the survey weight times a key size variable.

For sub-annual surveys, sub-annual rates as well as annual average rates (by summing the numerators and denominators of the sub-annual rates) for the preceding calendar year are to be reported. For other surveys, rates for the each round of the survey are to be reported.

Divisions are also required to provide descriptive information with these rates. The description should specify the unit for which the response rate has been calculated, as well as details on how the weighting was performed. If weighting is done by a key variable, the variable must be defined. Additionally, managers should try to define the invisible line between what is a partial response and a nonresponse in their survey. The reporting period covered by each response rate should be noted. Any deviations from the recommended guidelines should also be noted.

The response rates reported to the IMDB will form an integral part of the survey's Bi-annual Program Report (BPR). Augmenting this basic information with information from a longitudinal database maintained by the survey would be desirable.

FIGURE 1: RESPONDENT / NONRESPONDENT COMPONENTS AT THE DATA COLLECTION PHASE

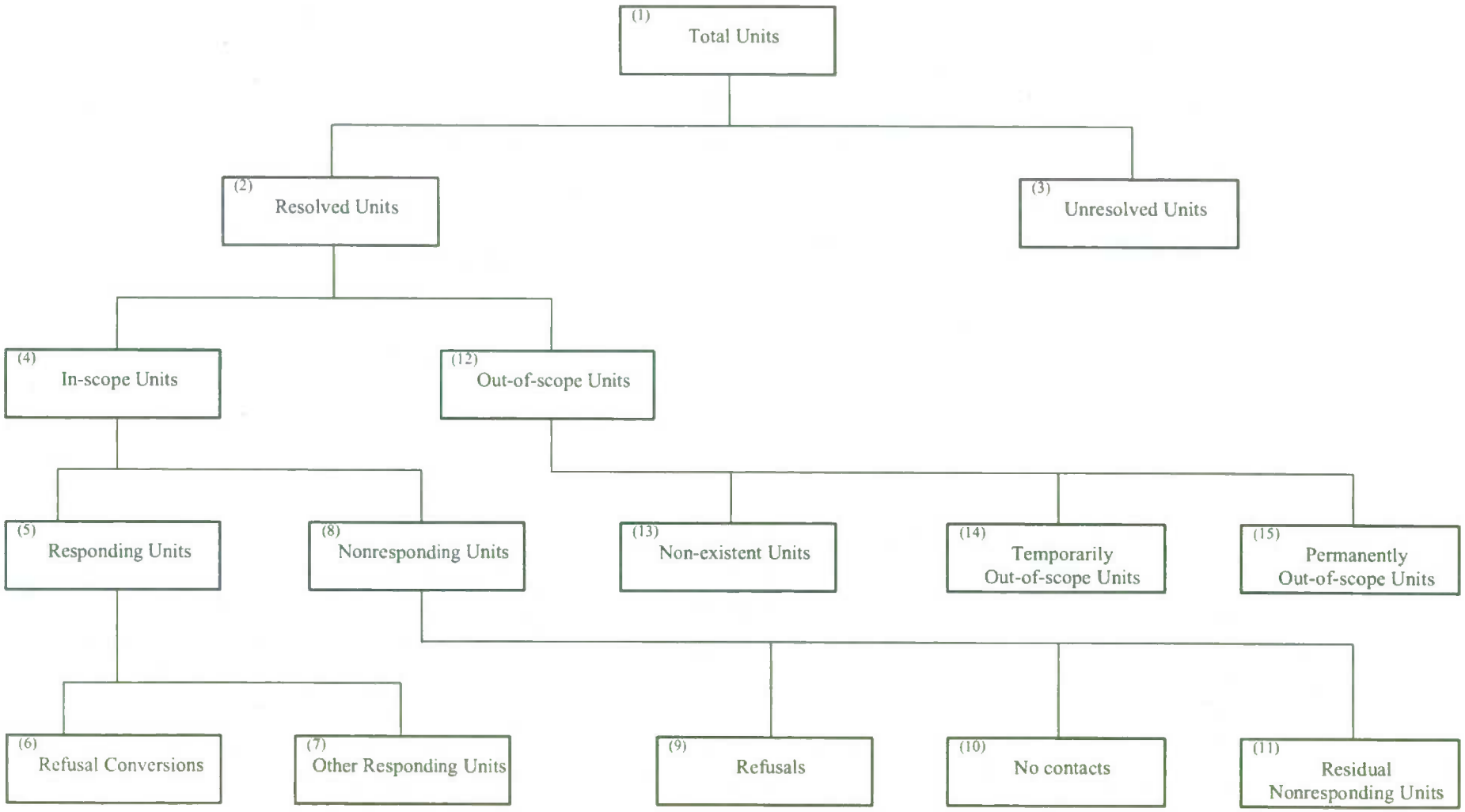
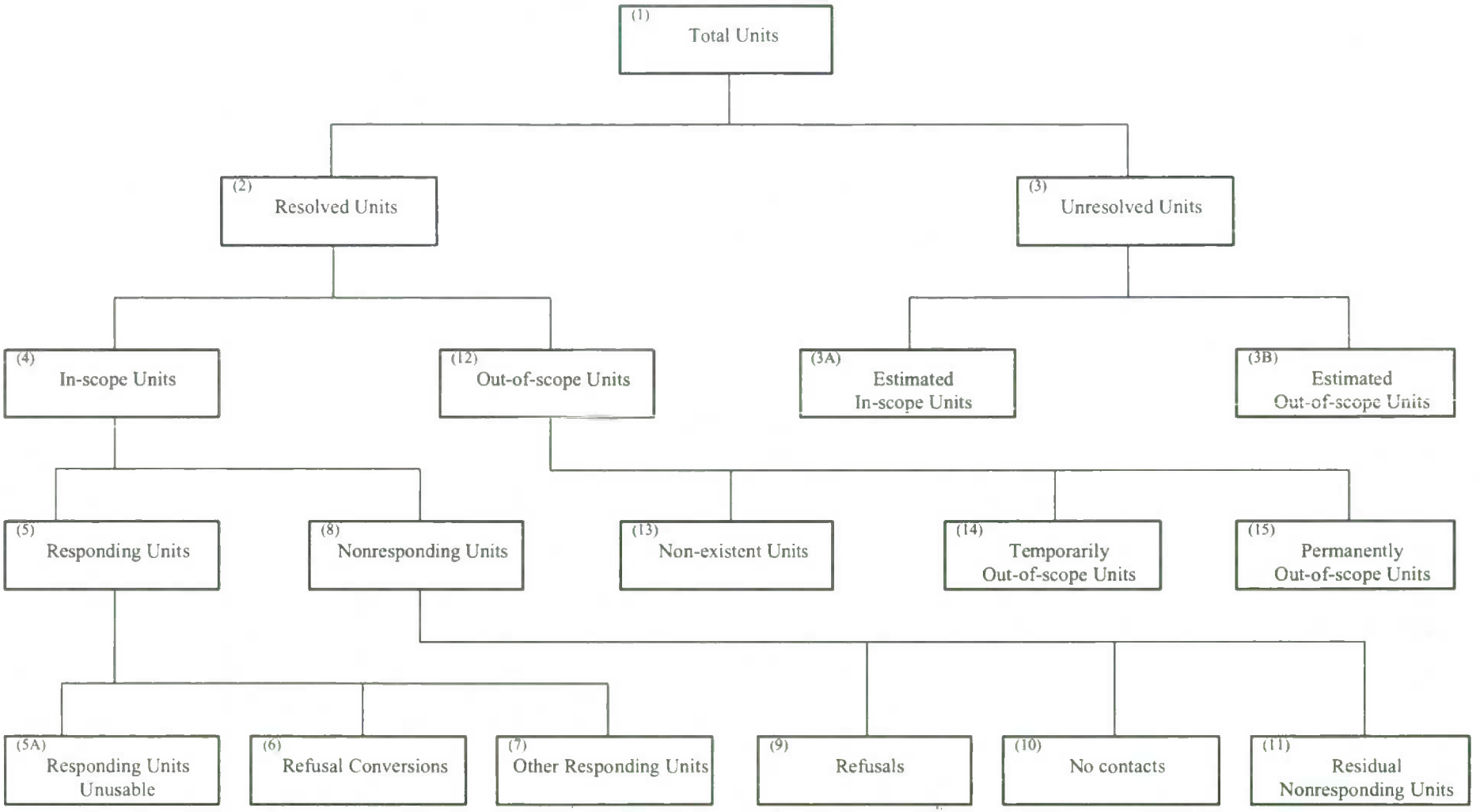


FIGURE 2: RESPONDENT / NONRESPONDENT COMPONENTS AT THE ESTIMATION PHASE



APPENDIX: Calculation of the Response Rate at the Estimation Phase

Total Number of Units (Box 1)	(1)					
Resolved Units						
In-Scope Units						
Responding Units						
Number of Unusable Responding Units (Box 5A)	(5A)					
Number of Refusal Conversions (Box 6)	(6)	+				
Number of Other Responding Units (Box 7)	(7)	+				
Total Number of Responding Units (Box 5 = Box 5A + 6 + 7)	(5)	=	=			
Nonresponding Units						
Number of Refusals (Box 9)	(9)					
Number of No Contacts (Box 10)	(10)	+				
Number of Residual Nonresponding Units (Box 11)	(11)	+				
Total Number of Nonresponding Units (Box 8 = Box 9 + 10 + 11)	(8)	=	=			
Number of In-Scope Units (Box 4 = Box 5 + 8)	(4)		=	=		
Out-of-Scope Units						
Number of Non-Existent Units (Box 13)	(13)					
Number of Temporary Out-of-Scope Units (Box 14)	(14)	+				
Number of Permanently Out-of-Scope Units (Box 15)	(15)	+				
Number of Out-of-Scope Units (Box 12 = Box 13 + 14 + 15)	(12)	=	→	+		
Number of Resolved Units (Box 2 = Box 4 + 12)	(2)			=	=	
Unresolved Units						
Estimated Number of In-Scope Unresolved Units (Box 3A)	(3A)					
Estimated Number of Out-of-Scope Unresolved Units (Box 3B)	(3B)	+				
Number of Unresolved Units (Box 3 = Box 3A + 3B)	(3)	=	→		+	
Total Number of Units (Verify that Box 1 from above = Box 2 + 3)	(1)				=	
Response Rate at the Estimation Phase						
Total Number of Responding Units (Box 5 from above)	(5)					
Number of Unusable Responding Units (Box 5A from above)	(5A)	-				
Responding Units at the Estimation Phase (Box 5 minus Box 5A)	(16)	=	=			
Number of In-Scope Units (Box 4 from above)	(4)					
Estimated No. of In-Scope Unresolved Units (Box 3A from above)	(3A)	+				
In-Scope Units at the Estimation Phase (Box 4 plus Box 3A)	(17)	=	+			
Response Rate at the Estimation Phase (Line 16 divided by Line 17) x 100%	(18)		=	x 100%	=	