



SAMPLE MAINTENANCE MANUAL FOR ECONOMIC STATISTICS

The Business Sampling Frame (BSF) consists of enterprises with different activities that have to be maintained on a regular basis to ensure a high quality frame. Administrative sources are used to update the database. The BSF database serves as a sampling frame for all surveys conducted by Stats SA. Samples are drawn in July/August every year. Every survey is required to follow Current Best Methods (CBM) to maintain their samples. This manual describes the sample maintenance procedures within Economic Statistics at Stats SA.

**This document was compiled by the division: Quality and Integration
Unit: Quality and Methodology.**

25 February 2003

1. The purposes of this manual

Firstly, this is a reference document for staff in their day-to-day work in the maintenance of the sample. This manual will be used in the process to maintain the sample.

Secondly, the manual will be a part of the support function that Quality and Methodology should play within Economic Statistics.

Thirdly, it is a part of the documentation for survey areas within Economic Statistics whose sampling frame is the Business Sampling Frame (BSF). Survey areas use the BSF for drawing samples. Every survey will have a separate sample allocated for their use and thus the maintenance should be done accordingly.

Fourthly, Stats SA is committed to quality work and quality management according to Systematic Quality Management (SQM) principles. In this context, the manual serves as a tool for standardisation and stabilisation of a part of the sample maintenance process. Standardisation and stabilisation are necessary for continuous quality improvement through improvements of the sample maintenance processes. **Continuous process improvements** mean that this manual is a “working” document that is updated at the pace of the process improvements. It will be revised and updated before the samples are drawn in June every year.

2. Responsibilities within Economic Statistics

Firstly, the maintenance of the BSF will be the responsibility of the component: Business Classification and Maintenance (BCM). All changes will be dealt with within this section. The BCM component will implement continuous process improvements and strive to deliver the best quality in terms of:

- Providing a directory from which mailing lists can be assembled for the dispatch of questionnaires
- Providing a population of businesses from which efficient sampling processes can be carried out and monitored
- Providing the basis for grossing-up results from sample surveys to produce population estimates
- Helping to prevent duplications and omissions in the collection of information on enterprises
- Improving congruence between the results of different surveys
- Helping to improve coverage or reveal inaccuracies
- Allowing coordination between the components conducting surveys. In cases where a frame is central and covers units from all the sectors of activity.

Business Sampling Frames are of fundamental importance to the compilation of economic statistics. Their coverage, comprehensiveness and quality vary considerably. A statistical Business Sampling Frame can be considered as a system of transforming data from administrative sources into data suitable for statistical use. In other words, Business Sampling Frames are designed to function as a bridge between administrative and statistical units.

The main users of Business Sampling Frames data are business survey areas. Taking into account that business statistics should both observe and describe a country's total productive (market) activity, the output of the ideal Business Sampling Frame can be defined as an up to date file of all statistical units active within the country's territory which generates value-added attributes, as well as their relevant statistical and administrative attributes. There exist new requirements for information on the structure of enterprises, e.g. data on takeovers, mergers and acquisitions. Additionally there has, for many years, been a rapid growth in the globalisation of enterprises– with the formation of multi-national companies spanning country boundaries. Business Sampling Frames also provide an element in reconciling requirements for increased information about enterprises and lightening their statistical response burden.

Secondly, the Survey Management System (SMS) will be used to indicate changes e.g. misclassification, mergers, deaths, etc.– thus a difference from the real-life situation for a specific enterprise. The accuracy, in terms of a standard error, in the estimates of change over time increases with a large overlap hence a large overlap in combination with survey feedback to the BSF is not recommended. The enterprises in the BSF, which are included in samples, will then be updated while those enterprises which are only included in the BSF (but not included in the samples) will not. This means that a large overlap in combination with survey feedback leads to biased estimates. From this point of view it may be advisable to update large enterprises with survey feedback because they are completely enumerated. In general, information from the surveys should rather be used as indicators in the maintenance processes connected to the BSF. It is also possible to use information on classification variables collected in the surveys as domains.

Thirdly, the survey areas will deal with their own samples and maintain their own samples for the specific year. A particular survey's information could be used for another survey but the purpose of the data collection and the level of presentation should be clearly understood. Information like addresses, respondent information could assist survey areas through the use of the centralised SMS, in the process of finding more correct information. Each survey area in collaboration with Quality and Methodology will be responsible to do sample maintenance through the use of the manual provided.

The classification in the BSF, in terms of economic activity, could sometimes be incorrect. In the majority of the economic surveys the enterprises included in the samples are asked, in the questionnaires, about their main activity. This information makes it possible for the survey areas to classify the enterprises in the samples into the correct economic activity thus making it possible to use the information from the surveys on correct economic activity as domains in the analysis of the results.

Currently, at Stats SA, the survey areas consider the incorrectly classified enterprises as:

- Non-response which are then imputed with the average in the stratum; or
- They use the incorrectly classified enterprises as correctly classified and they handle them like all the other enterprises in the industry.

The non-response method is based on the assumption that over-coverage and under-coverage even themselves out, so imputing the over-coverage is a way to compensate for the under-coverage. This is because there are, of course, enterprises incorrectly classified into other industries which should be correctly classified into the industry where the over-coverage was found. The method of keeping the incorrectly classified enterprises in the original industry and handling them as the other enterprises in the industry is also based on the assumption that over-coverage and under-coverage evens itself out.

The method of imputing an incorrectly classified enterprise with the average in the stratum is often preferable if the enterprise does not fit in the industry. For example, if a manufacturing enterprise is incorrectly classified into retail trade, it is likely that it is impossible to keep the original information from that enterprise in retail trade.

But, with both methods, correct information on economic activity is thrown away. It is possible to use the information on economic activity collected in the questionnaires as domains. This information could be used in the same way as all the other variables collected in the questionnaires, like gender, province etc. *If the correct information on economic activity is used as domains, the accuracy – in terms of coverage – in the point estimates will increase. But everything has a price and in this case the price will be larger variances. This is due to the fact that the domains based on the correct information on economic activity will cut through the original strata based on the information from the BSF.*

The method of using the correct information on economic activity as domains is straightforward if the survey covers the whole economy. For example, let us say that the survey covers all industries except agriculture– then it is impossible, with information from the survey, to find enterprises incorrectly classified into agriculture. In this instance this means that it would not be correct to use information from the survey to classify enterprises into agriculture. Conversely it would be impossible to use information from the survey to classify enterprises from agriculture into other activities. This would lead to under-coverage in the industries covered by the survey. If the correct economic activity for an enterprise is not covered by the survey, then it is recommended to use one of the methods currently in use e.g. to impute the enterprise with the average in the stratum or just to keep it in the incorrect activity.

Whether to use the correct information on economic activity as domains or not is of course a balance between smaller coverage error and larger variance – thus a balance between accuracy in the point estimates and larger variance. In practice there are probably a few important incorrectly classified enterprises having a large impact on the coverage error. And there are, as well, a lot of incorrectly classified

enterprises having a small impact on the coverage error just contributing to a larger variance. One way to take this balance into account is to combine the methods for handling incorrectly classified enterprises. By Creating a new variable, which is to be used as information on domain activity. For the majority of the enterprises in the sample the domain activity is the same as the stratum activity from the BSF, while for important incorrectly classified enterprises the domain activity is the information collected from the survey. Use this new variable as information on economic activity when the result is to be presented. Note that the estimation program must be able to handle domains cutting through strata in order to have correct variance estimates.

**STATISTICS SOUTH AFRICA SAMPLE MAINTENANCE
FOR ECONOMIC STATISTICS**

SAMPLE MAINTENANCE MANUAL

TABLE OF CONTENTS	Page
1. <i>The purposes of this manual</i>	2
2. <i>Responsibilities within Economic Statistics</i>	2
CHAPTER 1	8
GUIDELINES FOR USERS	8
1.1 <i>Introduction</i>	8
1.2 <i>This manual</i>	8
a. Maintenance of this manual	9
b. Exceptions /variations to the procedures.....	9
1.3 <i>Terminology</i>	10
a. Terminology used in business surveys.....	10
b. Other important terms	13
1.4 <i>The Unit Status Indicator (USI)</i>	13
SUMMARY TABLE OF COLLECTION PROCEDURES IN CHAPTER 2	17
CHAPTER 2	24
DATA COLLECTION PROCEDURES	24
2.1 <i>Introduction</i>	24
2.2 <i>Discussions of procedures in sample maintenance</i>	25
(1) SAMPLED UNIT HAS BEEN SOLD TO, MERGED WITH, TAKEN OVER BY, SPLIT OFF OR BROKEN UP – EVIDENT DOCUMENTATION AVAILABLE (USI 10).....	25
(2) CLOSED DOWN/LIQUIDATED DURING THE SURVEY PERIOD – EVIDENT DOCUMENTATION AVAILABLE (USI 11).....	37
(3) FINALLY LIQUIDATED – EVIDENT DOCUMENTATION AVAILABLE (BEFORE THE SAMPLE WAS DRAWN) (USI 09)	41
(6) UNDELIVERED QUESTIONNAIRE (USI 01).....	44
(9) DUPLICATE – DIFFERENT SAMPLED UNITS, BUT SAME BUSINESS (USI 28)	49
(11) INFORMATION OF ONE SAMPLED UNIT INCLUDED IN THE INFORMATION OF ANOTHER UNIT (OUTSIDE THE SAMPLE) (USI 13)	52
(12) FOREIGN BUSINESSES, E.G. BRITISH AIRWAYS, AIR FRANCE. THE FINANCIAL INFORMATION IS NOT AVAILABLE FOR THE SOUTH AFRICAN BUSINESS (USI 15)	55
(13) RECEIVED COMPLETED QUESTIONNAIRE FROM RESPONDENT BUT RE-CLASSIFIED – WITHIN THE SAMPLE - DOMAIN VARIABLE (USI 17)	55
(14) RECEIVED COMPLETED QUESTIONNAIRE FROM RESPONDENT BUT RE-CLASSIFIED – OUTSIDE THE SAMPLE (USI 18)	57
(15) SAMPLED UNIT UNABLE TO SUPPLY INFORMATION ON AN ENTERPRISE LEVEL (USI 05).....	58
2.3 <i>Flowcharts of procedures in the sample maintenance in Economic Statistics</i> ...	62
CHAPTER 3	83
ESTIMATION	83
3.1 <i>Integration of SMP into surveys</i>	83

3.2 USI facilities.....	83
3.3 Non-response.....	83
3.4 Imputation.....	84
3.5 Estimation.....	87
3.6 Outliers.....	87
3.7 Other systems issues.....	88
CHAPTER 4.....	89
GENERAL COMMENTS	89
4.1 Introduction.....	89
4.2 Justification for the current data collection procedures	90
4.3 Minimum procedures.....	92
4.4 Response rates.....	93
CHAPTER 5.....	95
GLOSSARY	95
1. Definitions.....	95
2. Acronyms.....	98
3. References.....	98

CHAPTER 1

GUIDELINES FOR USERS

1.1 Introduction

1. A business in the real world is characterised by its entity name(s), trading name(s), location address (es), business size, industry or industries, and its type of ownership. Some or all of these characteristics may change as businesses go through their life cycles.

2. As it is not realistic to expect the Business Sampling Frame (or any other lists of businesses used to select survey samples) to keep up with all the changes in the real world as they happen, it is inevitable that the representation of sampling units may deviate from corresponding real world business characteristics. Some of these differences, if not correctly treated by the survey areas, will result in large errors in survey estimates. Sample Maintenance Procedures (SMP) have been introduced to minimise these errors and to ensure that these situations are treated correctly and consistently across the entire range of Statistics South Africa (Stats SA) economic surveys.

3. The Unit Status Indicator (USI) is used to summarise the status of each unit selected in a business survey with regard to the uniform procedures. Its main purpose is to identify units which require special treatment in terms of data collection, estimation, and future dispatch in a manner that is consistent across all Stats SA business surveys. This enables consistent and coherent estimates to be produced by Stats SA economic surveys.

4. The Sample Maintenance Procedures Manual documents the procedures that should be followed when a difference is found between a business in the real world and its representation on the frame. This manual is intended to be used by supervisors, experienced editors and survey support staff (e.g. Quality and Methodology) in when they conduct of economic surveys.

5. The first chapter of this manual provides a background on the maintenance of the sample and exceptions from/variations to the procedures set out in this manual and whom this manual is for. It also defines the terminology used in this manual and provides an overview of the data collection procedures and the components of the USI. The second chapter explains the data collection procedures in detail. Other chapters provide additional information for the survey areas and their support staff, including justification of certain rules and procedures, some systems documentation, and operational tools for staff dealing with respondents on a day-to-day basis.

1.2 This manual

This manual is intended to be a reference tool for the use of supervisors, experienced editors and survey support staff (e.g. Quality and Methodology).

In general, all Stats SA surveys of businesses are covered by the Sample Maintenance Procedures. However, the extent to which procedures apply will depend on the source of the frame and the type of the sampling unit used by the survey. Currently all economic surveys use the enterprise as sampling unit.

This manual should be used as a first point of reference if a particular procedure requires clarification or if a problem is encountered. If the manual doesn't provide the required clarification then a query should be submitted to Quality and Methodology. Once queries are submitted to the Quality and Methodology staff member responsible for the support of all surveys, depending on the urgency of the query, the problem will be solved in collaboration with the specific subject matter specialist. Quality and Methodology staff is available to discuss problem cases with all survey areas if the need arises.

a. Maintenance of this manual

Businesses in the real world are not static and therefore a sample has to be maintained after selection. Selected businesses cease to operate, they merge, one is taken over by another, etc. There are also new businesses born after a sample was selected. This information is important for estimation, for future dispatch of questionnaires and for maintaining the Business Sampling Frame (BSF) of Stats SA.

The maintenance of this manual will be an ongoing review of the procedures. However, it will be updated and revised only once a year, in June, because that is when samples will be drawn within Economic Statistics. For errors found during the year the following review mechanism will be adopted:

- i. Quality and Methodology will review the affected procedure(s) and provide the survey with the interim treatment for the affected cases. The interim treatment will be documented in the manual under "Review of the Manual" category.
- ii. If appropriate (based on the magnitude of the errors and the changes required), the modified procedures will be circulated for comment and submitted for approval by all survey areas and Quality and Methodology.
- iii. Once the changes/corrections are finalised they will be circulated to the stakeholders.
- iv. The modified procedures will be documented in the electronic version of this manual, while the actual use of the updated and revised manual will be undertaken in June every year. Commonly occurring situations not covered by this manual will be treated in the same way as errors in procedures.

b. Exceptions /variations to the procedures

Those specific to a survey: the exceptions to or variations of the procedures that are specific to suit a particular survey have to be approved by Quality and Methodology. If Quality and Methodology has approved the variations they will be included in this

manual. The following process should be followed if an exception or a variation is required:

- i. The requirement for an exception/variation should be discussed with the Quality and Methodology staff responsible for the support of the survey.
- ii. The justification for the exception should be documented by the survey area and submitted to Quality and Methodology under the Exceptions/Variations category and e-mailed to the Quality and Methodology support team. Statistical Support recommendations by the Quality and Methodology staff, as to whether the exception/variation is supported should be provided together with any additional documentation.
- iii. The Quality and Methodology team will review the documentation. Additional documentation from the survey area may be required at this stage.
- iv. The Quality and Methodology team will provide a recommended treatment.

Exceptions or variations that are not specific to the survey (e.g. queries relating to a specific unit or mistake found in the procedures) should be submitted as a general Quality and Methodology query.

It is impossible to predict every possible scenario that can occur in a survey. It is therefore anticipated that there will be a small number of units for which standard treatment is not appropriate although the general principles will apply. These units are likely to be medium or large businesses with large turnovers and have complex structures. The procedures for these units will be devised over time as the actual cases occur and as the knowledge base is built up on the database. If an exceptional case occurs the details should be submitted as a query, unless a similar situation has already been raised and resolved.

1.3 Terminology

This section articulates the terminology used in business surveys when describing units on the frame and in the sample and their relationship to the real world businesses. The rest of this manual relies on the reader's understanding of the terms and ability to interpret these terms in the context of the specific survey. In addition, this section provides definitions of other terms, which are considered important for the understanding of this manual but may not be in general use in the survey areas.

a. Terminology used in business surveys

Statistical unit

Is a unit about which statistics are tabulated, compiled or published. The statistical units are created from and linked to the South African Revenue Services (SARS) administrative data. The entities and their characteristics shown are those that are visible to the BSF user through the BSF system.

Enterprise (EN)

Characteristic	Purpose	Possible values
Enterprise number	Unique identifier	'Enxxxxxxxx' where 'x' is a digit
Legal name	Name of the legal unit	
Total number of employees	Size measure	
Turnover	Size measure	Calculated from VAT
Sector	Economic sector	'1'=Agriculture, hunting and related services '2'=Mining and quarrying '3'=Manufacturing '4'=Electricity, gas and water supply '5'=Construction '6'=Wholesale and retail trade; repair of motor vehicles, motor cycles and personal and household goods; hotels and restaurants '7'=transport, storage and communication '8'=Financial intermediation, insurance, real estate and business services '9'=Community, social and personal services
Sic	Industrial activity according to SIC v5, 1993	See SIC V 5, 1993
Owner type	Owner type	'1'=Individual '2'=Partnership '3'=Public Company Ltd/Bpk '4'=Pty Ltd/ Edms Bpk '5'=Public Corporations '6'=Close Corporations CC/BK '7'=Co-operative Societies '8'=Local Authorities '9'=Government '0'=Other - churches etc.
Life cycle		'birt'= Birthed 'ceas'=Ceased 'reac'=Reactivated 'inva'=Invalid 'activ'=Activated
Random number		A number between 0 and 1

Statistical unit structure

- The statistical unit "enterprise" (**EN**) is a legal unit – (a LU corresponding to a company or income tax unit) – or the smallest combination of LUs that includes and, directly controls all the functions necessary to carry out its production activities.
- A Kind-of-Activity Unit (**KAU**) engages in one or predominantly one kind of economic activity without being restricted to a geographic location.
- A Geographic Unit (**GEO**) is normally an unbroken physical area/site on or from which one or predominantly one kind of economic activity takes place on a permanent basis.

Sampling unit

A sampling unit is a statistical unit on the frame, which is available for selection. A sampling unit is characterised by a unique identifier (e.g. a Enterprise Number [EN] on the Business Sampling Frame), a unit identification name (e.g. legal entity name) and attributes describing the unit (e.g. turnover, industry type).

Selected sampling unit (Selected unit)

Sampling unit that is selected in the sample.

Sampling unit unique identifier

A unique identification number which for all Business Sampling Frame based surveys is an Enterprise Number (EN).

The unique identifier are statistical concepts defined for the purpose of collecting data, whilst the unit identification name and attributes are real world concepts which describe the business in the real world. In theory, on a perfect frame, the identification name and attributes of the unit would be identical to their real world counterparts; however, often in practice, the frame identification name and attributes are different from those in the real world. Following are the terms necessary to describe the differences between identification features recorded on the frame and their real world counterparts.

Attributes

These are the business activity, company structure, trading name, liabilities, and turnover. A description of a unit's attributes on the frame will be referred to as frame attributes, whilst the real world counterpart will be referred to as either real world attributes or business operations.

b. Other important terms

Sampling Frame

The sampling frame is a list of all units in the survey population available for selection at the time of sample selection.

Observation unit

An observation unit is a unit for which the data that does not equate to the statistical unit on the frame are reported, i.e. it is an aggregation or dissection of statistical units.

1.4 The Unit Status Indicator (USI)

The Unit Status Indicator (USI) summarises the status of each unit selected in a business survey with regard to the uniform procedures. Its main purpose is to identify units which require special treatment in terms of data collection, estimation, future dispatch and frame feedback (only use this type of feedback as an indicator for the SMS) in a manner that is consistent across all Stats SA business surveys.

The USI may be thought of as either a collection of separate flags or as a composite flag.

To maintain the sample, each selected sampling unit has a unit status indicator code (USI code) to indicate its status. For example, the initial USI code of a selected unit is '00' and this code will change to '11' when the unit closes down. A unit status indicator's main purpose is to identify selected units that require special treatment in terms of:

- data collection;
- estimation;
- future dispatch; and
- feedback to the BSF

in a manner that is consistent across all Stats SA business surveys.

After questionnaires are dispatched to all the selected units, these units have to respond by sending back their completed questionnaires within a specified time period. Information received from a unit that responded is used to determine its status. If a unit does not respond then information is obtained from other sources (e.g. Brabys) to determine its status.

The unit statuses as used by the business surveys are reflected in Table 1 below.

Table 1 - Unit Status Indicator (USI) codes for the 2002 surveys to be used in the Survey Management System (SMS) of Stats SA

USI code	Description
00	Outstanding questionnaire
01	Undelivered questionnaire
02	Undelivered reminder
03	Enterprise under investigation
04	Untraceable – Process of investigation finalised
05	Sampled unit unable to supply information on an enterprise level ¹
06	Extension after due date
07	Closed down – Evident documentation available (before the sample was drawn)
08	Dormant – Evident documentation available
09	Finally liquidated – Evident documentation available ² (before the sample was drawn)
10	Sampled unit has been sold to, merged with, taken over by, split-off or broken up – Evident documentation available
11	Closed down/Liquidated during the survey period – Evident documentation available
12	Information of one sampled unit included in the information of another sampled unit (within the sample)
13	Information of one sampled unit included in the information of another unit (outside the sample)
14	Received completed questionnaire from respondent with no employees working in the sampled unit, e.g. Property owning and letting companies (SIC 84110)
15	Foreign businesses, e.g. British Airways, Air France. The financial information is not available for the South African business ³

¹ Create observation unit on the SMS.

² If still in the process of liquidation the completed questionnaire must be supplied to Stats SA. The USI code will reflect '00' until receipt of the completed questionnaire whereafter the USI code will reflect '22'.

USI code	Description
16	Selected unit situated outside the borders of South Africa ⁴
17	Received completed questionnaire from respondent but re-classified – Within the sample - domain variable
18	Received completed questionnaire from respondent but re-classified – Outside the sample ⁴
19	Only a Part 1 or a Part 2 questionnaire received, respondent to be followed-up (SEOGR survey)
20	Legal action to be taken
21	Manual imputation
22	Received completed questionnaire from respondent
23	Received financial statement or pay-roll printouts from respondent
24	Omit the sampled unit for a short period, e.g. computer systems currently not working, a fire broke out in the factory, etc.
25	Checked questionnaire
26	Edited questionnaire
27	Edited financial statement
28	Duplicate – different sampled units, but same business
29	ASO refers investigation to supervisor

Note – Changes to name, address and other contact details are to be changed on the SMS. Before any name and/or address changes are done the collection clerk should phone the respondent and enquire if this name and/or address change did not occur because of a demographic event. A demographic event includes that of sold, merged, taken-over, split-off and break-up i.e. USI code 10. If not a demographic event then the collection clerk must allocate the applicable USI code according to the new information received from the respondent.

³ A questionnaire must not be sent in the next survey period. Only applicable to Financial Statistics.

⁴ Out of scope.

If the only change is a name change then during the next round of surveying the questionnaire must be posted with the new name and/or address. It is important to remember that a USI code cannot be allocated for a change in name and/or address.

These USI codes will need to have their values set by clerical action, except in cases where the values are carried through from another survey or earlier processing cycle. USI codes are derived automatically from previous SMS survey cycles where no future dispatch should happen whilst estimation codes are derived automatically from response codes indicated by the USI codes for a specific period. The person responsible for this action is the Survey Statistician. During the estimation phase of the survey cycle, the USI codes are used to ensure each unit is treated correctly in the production of estimates and variances. The relevant codes are set prior to the survey's dispatch to ensure that respondents will not have unnecessary contact with Stats SA, thus ensuring that respondent relationships are maintained.

The actions indicated by the USI code values are not tied to any particular phase of the processing cycle since they have no effect on the survey's outputs. For example, the USI code is used in the carrying forward of response codes to the next survey cycle. The USI code in terms of (i.t.o.) responses look "at the past". In the end of a processing cycle these codes show how a unit responded and the measures that were taken during that cycle. The USI code i.t.o. the next survey cycle looks "to the future" in that it determines whether or not the unit may be dispatched in the future.

The USI code flags the dispatch action for the unit in later cycles, as well as the responses through the life cycle of the survey. The USI code indicates how a unit is to be treated during estimation. This will be done through the use of SAS procedures, in the process of imputation and estimation.

SUMMARY TABLE OF COLLECTION PROCEDURES IN CHAPTER 2

USI Code	Description	Action	Contribution to survey estimates	Further dispatch action	Response
00	Outstanding questionnaire	Respondent didn't respond. Apply applicable USI code.	Impute	Dispatch	No
01	Undelivered questionnaire	Follow minimum 'trace' procedures if no information was found then treat as untraceable.	Impute	No dispatch	No
02	Undelivered reminder	Apply applicable USI code according to new information received.	Impute	No dispatch	No
03	Enterprise under investigation	Apply applicable USI code according to new information received.	Impute	No dispatch	No
04	Untraceable – Process of investigation finalised	If minimum 'trace' procedures were followed and no information were found then treat as untraceable.	Impute	No dispatch	No

05	Sampled unit unable to supply information on an enterprise level	Unit cannot provide information on the structural level (enterprise level). Create observational units on the SMS.	Value/Impute	Dispatch to the observational units.	No
06	Extension after due date	Apply applicable USI code according to new information received.	Impute	Dispatch	No
07	Closed down – evident documentation available (before the sample was drawn)	Unit indicates that operations closed down and no future income will be generated. Assign USI 07.	Adjust weights for non-response	No dispatch	No
08	Dormant – Evident documentation available	Unit indicates that operation is dormant; proof that the enterprise is finally closed down should be collected. Assign USI 08.	Adjust weights for non-response	No dispatch	No
09	Finally liquidated – Evident documentation available (before the sample was drawn)	Unit indicates ceased operations, liquidated. Assign USI 09.	Adjust weights for non-response	No dispatch	No

10	Sampled unit has been sold to, merged with, taken over by, split off or broken up – Evident documentation available	Both enterprises are included in the sample then assign applicable USI code.	<i>Enterprise A:</i> Capture as zero's and <i>Enterprise B:</i> Σ (Enterprise A + Enterprise B)	<i>Enterprise A:</i> No dispatch. <i>Enterprise B:</i> Dispatch and capture the relevant information (e.g. enterprise name, address, contact information etc.).		No
		Only one enterprise is included in the sample.	Large effect on the estimates	Enterprise A: Impute	Enterprise A: No dispatch	No
		Enterprise A is included in the sample and enterprise B is not included in the sample but on the sampling frame). Assign applicable USI code.	Small effect on the estimates	<i>New Enterprise:</i> Value	<i>New Enterprise:</i> Dispatch and capture the relevant information (e.g. enterprise name, address, contact information etc.)	No
11	Closed down/Liquidated during the survey period – Evident documentation available	Unit indicates ceased operations (e.g. closed down /liquidation). Assign applicable USI code.	Impute	No dispatch		No
12	Information of one sampled unit included in the information of another sampled unit (within the sample)	Primary unit included then treat as response. Assign applicable USI code.	Value	Dispatch		No
		Secondary unit included. Assign USI 12.	Zero	No dispatch		

13	Information of one sampled unit included in the information of another unit (outside the sample)	Unit unable to respond because the unit's information is included in another unit that falls outside the sample. Assign USI 13	Adjust weights for non-response	No dispatch	No
14	Received completed questionnaire from respondent with no employees working in the sampled unit, e.g. Property owning and letting companies (SIC 84110)	Received completed questionnaire. Capture information received.	Value	No dispatch	Yes
15	Foreign businesses, e.g. British Airways, Air France. The financial information is not available for the South African business	Unit situated within the borders of South Africa but the payrolls are kept overseas. Assign USI 15.	Impute	No dispatch	No
16	Selected unit situated outside the borders of South Africa	Unit situated outside the borders of South Africa and payrolls are kept overseas. Assign USI 16.	Impute	No dispatch	No

17	Received completed questionnaire from respondent but re-classified – Within the sample - domain variable	Economy-wide surveys. Capture information received.	Value	Dispatch	Yes
		Sector oriented surveys. Capture information received.	Value	Dispatch	
18	Received completed questionnaire from respondent but re-classified – Outside the sample	Economy-wide surveys. Assign appropriate USI code.	Value, Impute or Adjust weights for non-response	Dispatch/ No dispatch	No
		Sector oriented surveys. Assign USI 18.	Adjust weights for non-response	No dispatch	
19	Only a Part 1 or a Part 2 questionnaire received, respondent to be followed-up (SEOGR survey)	Received part of the information. Capture information received and collect outstanding information.	Impute/Value	Dispatch	No
20	Legal action to be taken	Respondent refuse to provide information. Assign USI 20.	Impute	No dispatch	No
21	Manual imputation	Only applicable for survey areas that still impute manually. Assign USI 21.	Impute	Dispatch	No

22	Received completed questionnaire from respondent	Received information. Capture information received.	Value	Dispatch	Yes
23	Received financial statement or payroll printouts from respondent	Received information. Capture information received.	Value	Dispatch	Yes
24	Omit the sampled unit for a short period, e.g. computer systems currently not working, a fire broke out in the factory, etc.	Temporarily omitted unit. Capture the extension date and follow up after the date expired. Assign USI 24.	Impute	Dispatch	No
25	Checked questionnaire	Received information. Capture information received.	Value	Dispatch	Yes
26	Edited questionnaire	Received information. Capture information received.	Value	Dispatch	Yes
27	Edited financial statement	Received information. Capture information received.	Value	Dispatch	Yes
28	Duplicate – different sampled units, but same business	Primary duplicate. Treat as response. Capture information received.	Value	Dispatch	Yes
		Secondary duplicate. Assign USI 28	Zero	No dispatch	No

29	ASO refers investigation to supervisor	Apply applicable USI code according to new information received after the investigation is done.	Impute	No dispatch	No
----	--	--	--------	-------------	----

Note: For clerical use, USI codes 00, 01 02, 03, 06, 19, 20, 21 and 24 are treated as outstanding thus no information is received from respondent (non-response) These USI codes is for monitoring within every survey.

CHAPTER 2

DATA COLLECTION PROCEDURES

2.1 Introduction

This section discusses in detail most of the data collection procedures outlined in Chapter 1 of this manual. The discussion is structured for most of the data collection procedure as follows:

Respondent contact: defines criteria which the respondent has to meet in order to satisfy the definition of the procedure and provides instructions which should be followed if the criteria are not met.

Data processing: defines procedures for data processing for each SMP treatment.

SMS: defines information that needs to be updated on the SMS.

USI codes: defines the Unit Status Indicator (USI) codes, which apply to each SMP treatment. Estimation and future cycles: defines the estimation and action in later cycles of the survey if the sampling unit is selected in the following cycle of the survey. Contribution to the estimates can be either a value or zero. Value refers to reported or imputed data that contributes to the estimates (please note a value may be zero for some data items). Zero is used to indicate that a unit contributes zero to all estimates.

Discussions in this document refer to sampling unit and real world identification names and frame attributes and business operations (real world attributes) as defined in the terminology section. Therefore it is important, that in each specific example, users understand how these concepts relate to the specific cases.

All of the data collection procedures assume that if a difference between the sampling unit and a real world business is detected, survey staff will be able to determine the correct representation of the business and indicate the difference through the SMS by assigning the applicable USI. However, in a small proportion of cases, the differences between the sample and the real world are a result of the complex restructuring of the business and it is not easy to determine without specialised profiling knowledge how the unit should be represented on the SMS. In these cases, before SMP treatment can be decided upon, survey areas expert staff needs to consult profiling experts to determine the correct USI code to allocate.

When multiple changes happen to a unit and it is not clear which procedure applies (using this manual), survey staff should consult their statisticians to determine the final USI code.

In the next section attention will be focussed on the sample maintenance procedures – that is, structural changes within businesses such as the treatment of closed

down/dormant enterprises, name changes, take-overs, mergers, enterprises in liquidation and misclassified enterprises.

Weighting:

The principal methods for non-response adjustments are weight adjustment and imputation. Imputation will be used for item non-response and unit non-response. An effect from non-response is that the sampling error will increase due to the fact that the number of observations is reduced.

Consider a population with H strata containing N elements (sampling units).

Let

N_h = number of population elements in the h_{th} stratum

n = total sample size

n_h = sample size in the h_{th} stratum

w_h = weight in the h_{th} stratum

m_h = total number of responding enterprises in the h_{th} stratum

v_k = adjustment weight per strata for non-response

Where the selection probability of an enterprise in stratum h is n_h / N_h and the inverse of this selection probability is the design weight, $w_h = N_h / n_h$.

If there is no non-response, then the weight that would be used is $w_h = N_h / n_h$ for all enterprises that were sampled in stratum h . However, in reality we have to deal with non-response. The Sample Maintenance Manual outlines measures to be taken in attempting to reduce non-response. Survey areas should do everything in their power to collect information from respondents to ensure that non-response is reduced.

The adjustment weight per strata for non-response is, $v_k = n_h / m_h$.

The final weight, $w_k = (N_h / n_h) v_k = w_h \cdot v_k = N_h / m_h$ for enterprise k in stratum h and this weight is then used in the estimation.

2.2 Discussions of procedures in sample maintenance

(1) SAMPLED UNIT HAS BEEN SOLD TO, MERGED WITH, TAKEN OVER BY, SPLIT OFF OR BROKEN UP – EVIDENT DOCUMENTATION AVAILABLE (USI 10)

When an enterprise is sold, merged, taken over, split off or broken up, then a number of things can materialise:

Take over

Scenario 1:

Enterprise A (the smaller enterprise) has been taken over by enterprise B (the larger enterprise). The measure of size is turnover. If both enterprises (enterprise A and enterprise B) are included in the sample then enterprise A must be captured as zero's and enterprise B will be the sum of enterprise A and enterprise B. Enterprise B's weight will be used. The smaller enterprise and the largest enterprise would be determined by taking their turnover into account, as on the sampling frame. For example, if we want to estimate the variable 'employment' then firstly we would look at the turnover to determine the largest enterprise that will be used in the estimation. The result would be that the sum of enterprise A and enterprise B would be calculated by taking the sum of the employment of the enterprises A and B because this is the variable that we would like to estimate.

Scenario 2:

Large effect on the estimates

It could also happen that enterprise A (the smaller enterprise) is included in the sample and taken over by enterprise B while enterprise B (the larger enterprise) is not included in the sample. It should be noted that as the new enterprise name has changed as well as a change in its activity took place then enterprise A should be imputed for the remainder of the survey cycle.

Small effect on the estimates

If one enterprise takes over another enterprise, which is not in the sample then the effect is small and the results could be accepted. The small effect could be described as being less than 5% within a subgroup (3-digit), 1% within a group in terms of turnover as on the sampling frame. This would usually be the case if the larger enterprise (enterprise B) was in the sample and the take-over happened with the smaller enterprise with enterprise B has taking over enterprise A.

The following examples will illustrate the two stated scenarios

Example 1

Statistical units selected in a certain survey have the following sample attributes:

Enterprise A

Frame identification name (Legal name): MGX Customer Service Excellence (Pty) Ltd.
Trading name: CSX.

Mailing address: P.O. Box 1679
Bramley
South Africa
2018

Sampling unit unique identifier: EN 15356983.

Classification: SIC 96490 (Other recreational activities).

Turnover: R 49 574 781,77053.

and

Enterprise B

Frame identification name (Legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638
Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Classification: SIC 8899 (Other business activities n.e.c.).

Turnover: R 61 039 185,43131.

Mr. Phillips (the accountant of MGX Business Continuity Solutions Pty Ltd) phoned to say that he had received a survey questionnaire for MGX Business Continuity Solutions Pty Ltd as well as a questionnaire for CSX, but that CSX was taken over by MGX Business Continuity Solutions Pty Ltd and it would be impossible to provide information on the old structures of the enterprises. He wanted to know if he should complete the questionnaire on behalf of the 'new' enterprise namely MGX Business Continuity Solutions Pty Ltd. The processing clerk asked if MGX Business Continuity Solutions Pty Ltd **had taken over** the business from the company indicated on the survey questionnaire namely CSX. Mr. Phillips confirmed that this was the case. The processing clerk asked **when** MGX Business Continuity Solutions Pty Ltd had taken over CSX. Mr. Phillips said MGX Business Continuity Solutions Pty Ltd took over the business in September 2002 (during the reference period).

Application of scenario 1 (Both enterprises are included in the sample)

Enterprise B (MGX Business Continuity Solutions Pty Ltd) is the largest enterprise because its turnover is greater than enterprise A's (CSX) turnover, – the smaller enterprise.

The smaller enterprise and the largest enterprise would be determined by taking their relevant turnovers, as on the sampling frame.

Turnover_{Enterprise B} > Turnover_{Enterprise A}
 Turnover_{Enterprise B} of R 61 039 185,43131 > Turnover_{Enterprise A} of R 49 574 781,77053.

Both enterprises are included in the sample	
Description	Action
Enterprise A (CSX)	Captured as zeros
Enterprise B (MGX Business Continuity Solutions Pty Ltd)	Σ (Enterprise A + Enterprise B)

Use Enterprise B's weight.

If the take-over happened before the sample was drawn the treatment would be quite similar to the treatment of a take-over that happened during the reference period except that for the whole survey cycle the “new enterprise” information would be collected and the necessary name changes would be done on the SMS.

If it is possible to collect the information and the take over just happened then the respondent would still be able to provide Stats SA with the necessary information (because we ask information for the period t_{-1}). In this case if the respondent completed both questionnaires for enterprise A and enterprise B, and the following information was given for both enterprises, then the following is applicable.

Enterprise A (CSX) information received:

Employment: 36 employees
 Salaries and wages: R 90 000 for the quarter.

Enterprise B (MGX Business Continuity Solutions Pty Ltd) information received:

Employment: 46 employees
 Salaries and wages: R 150 000 for the quarter.

As explained above the larger enterprise is enterprise B and the smaller is enterprise A.

Capturing:

For Enterprise A (CSX) capture zeros for all variables in the questionnaire.
 For Enterprise B (MGX Business Continuity Solutions Pty Ltd) capture the following variables: employment and salaries and wages as follows:

Employment: **82 employees** (36 + 46).

Salaries and wages: **R 240 000 per quarter** (R 90 000 + R 150 000).

In the estimation part of it the weight of enterprise B (MGX Business Continuity Solutions Pty Ltd) will be used as calculated when the sample was drawn.

If the take over happened during quarter three, then quarter four's information would be collected from the new enterprise namely MGX Business Continuity Solutions Pty Ltd and a questionnaire would be dispatched in quarter four to the new enterprise (MGX Business Continuity Solutions Pty Ltd). When the new sample is drawn then hopefully the take-over would be recorded as it is observed in the real life situation and we would sample the newly-created enterprise.

In quarter four, information would be collected from the respondent and captured as indicated on the questionnaire using the weight of the old enterprise B e.g.

Employment: **92 employees.**

Salaries and wages: **R 300 000 per quarter.**

Use weight of enterprise B.

The action required is to:

- i. Inform the respondent that a completed questionnaire is required for the new enterprise namely MGX Business Continuity Solutions Pty Ltd (CSX and MGX Business Continuity Solutions Pty Ltd). Also, inform the respondent that a questionnaire for the old CSX and MGX Business Continuity Solutions Pty Ltd should not be completed in future, and that only one questionnaire containing information of the new enterprise (MGX Business Continuity Solutions Pty Ltd) after the take-over will be required;
- ii. Capture all relevant information required for future dispatch on the SMS; and
- iii. Assign a USI code of 10 to enterprise A and assign a USI code of 22 to enterprise B.

Example 2

Application of scenario 2 (Only one enterprise is included in the sample. Enterprise A is included in the sample and enterprise B is not included in the sample)

Enterprise B (MGX Business Continuity Solutions Pty Ltd) is the largest enterprise because it's turnover is greater than that of enterprise A's (CSX) turnover (the smaller enterprise). Enterprise A (the smaller enterprise) is included in the sample and is taken over by enterprise B (the larger enterprise) while enterprise B is not included in the sample. It should be noted that the new enterprise name has changed as well as that a change in its activity took place.

The smaller enterprise and the larger enterprise would be determined by taking their relevant turnovers, as on the sampling frame.

$$\text{Turnover}_{\text{Enterprise B}} > \text{Turnover}_{\text{Enterprise A}}$$

$$\text{Turnover}_{\text{Enterprise B}} \text{ of R } 61\,039\,185,43131 > \text{Turnover}_{\text{Enterprise A}} \text{ of R } 49\,574\,781,77053.$$

Only one enterprise is included in the sample. Enterprise A is included in the sample and enterprise B is not included in the sample	
Description	Action
Enterprise A (CSX)	Impute
Enterprise B (MGX Business Continuity Solutions Pty Ltd)	Not applicable because the enterprise is not forming part of the sample.

Use weight of enterprise A.

If the take over happened before the sample was drawn the treatment would be quite similar to the treatment of a take over that happened during the reference period except that for the whole survey cycle the information would be imputed for enterprise A and a USI 10 would be assigned.

If it is possible to collect the information and the take over just happened then the respondent would still be able to provide Stats SA with the necessary information (we ask information for the period t_{-1}). In this case, if the respondent completed the questionnaire for enterprise A and the information below was given for enterprise A the following action is applicable to this situation:

If the respondent could not complete the questionnaire because he/she could not break down the information in the format that we wanted him/her to, then enterprise A should be imputed for the whole of the survey period till new samples are drawn.

If the effect would be large on the estimates then the following would be applied:

Enterprise A (CSX) information received:

Employment: 36 employees
Salaries and wages: R 90 000 for the quarter.

Enterprise B (MGX Business Continuity Solutions Pty Ltd) no information received because not drawn in the sample for the specific survey.

Employment: unknown.
Salaries and wages: unknown.

As explained above the larger enterprise is enterprise B and the smaller is enterprise A.

Capturing:

For Enterprise A (CSX) impute all variables in the questionnaire.

For Enterprise B (MGX Business Continuity Solutions Pty Ltd – before take over) not relevant to specific survey (not part of the sample).

If the take over happened in quarter three, then quarter four's information would be imputed and a questionnaire wouldn't be dispatched in quarter four to the new enterprise (MGX Business Continuity Solutions Pty Ltd – after take over). When the new sample is drawn then hopefully the take over would be recorded as it is observed in the real life situation and we would sample the newly created enterprise.

In quarter four, information would be imputed for all variables in the survey for enterprise A, e.g.

Employment: **Impute.**

Salaries and wages: **Impute.**

Use weight of enterprise A.

If the effect would be small on the estimates then the following would be applied:

If an enterprise takes over another enterprise, which is not in the sample and the effect is small the results could be accepted.

Firstly, determine whether the effect is small or large. A small effect could be described as being less than 5% within a subgroup, 1% within a group. The turnover would be used to determine whether we could accept the results as received from the respondent. In this case the respondent did complete the questionnaire but used the information of the new enterprise namely MGX Business Continuity Solutions Pty Ltd. If the effect of the take over would be small in terms of the turnover then the results could be accepted and the information could be used as received from the respondent.

Enterprise A (CSX): no information received because the respondent completed the questionnaire using the newly created enterprise (MGX Business Continuity Solutions Pty Ltd) information:

Employment: unknown.

Salaries and wages: unknown.

Before the take over this would be the situation: From Enterprise B (MGX Business Continuity Solutions Pty Ltd) no information would be received because this enterprise wasn't drawn in the sample for the specific survey.

Employment: unknown.

Salaries and wages: unknown.

But after the take over the completed information received from the respondent would capture the information of both enterprises thus the new enterprise (MGX Business Continuity Solutions Pty Ltd):

Employment: 102 employees.
 Salaries and wages: R 360 000 per quarter.
 Use weight of enterprise A.

As explained above the larger enterprise is enterprise B and the smaller is enterprise A.

Taking the turnovers as on the sampling frame and then calculating the percentage contribution per subgroup and/or per group for the “old” and the ‘new’ enterprises would determine the effect. The turnover of the new enterprise should be collected from the respondent so that the necessary calculations could be done.

- Enterprise A’s (CSX) turnover is R 49 574 781,77053 on the sampling frame;
- Enterprise B’s (MGX Business Continuity Solutions Pty Ltd) turnover is R 61 039 185,43131 on the sampling frame; and
- Enterprise New’s (MGX Business Continuity Solutions Pty Ltd) turnover is R 55 134 555,12631 as received from the respondent.
- Total turnover of the subgroup is R 1 161 039 185,43131 and the total turnover within a group is R 2 361 039 185,43131.

Description	Less than 5% within a subgroup (3 digit)	Less than 1% within a group	Effect
Enterprise A (CSX)	4,3%	2,1%	Small
Enterprise B (MGX Business Continuity Solutions Pty Ltd) - <i>before take over</i>	5,3%	2,6%	
Enterprise New (MGX Business Continuity Solutions Pty Ltd) - <i>after take over</i>	4,8%	2,3%	

From the table above it can be seen that the effect on the estimates would be small. **Note:** that you can either use the turnover of the new enterprise (MGX Business Continuity Solutions Pty Ltd) received from the respondent, which fall into the category less than 5% within a subgroup contribution, or you could use the less than 1% within a group category to determine whether the new enterprise (MGX Business Continuity Solutions Pty Ltd) has a small effect or a large effect on the estimates. Enterprise New (MGX Business Continuity Solutions Pty Ltd) has an effect of 4,8% to the subgroup and thus it could be seen that the effect is small and this situation does fit into this scenario.

Capturing:

For Enterprise A (CSX) capture all variables in the questionnaire as zero.

For Enterprise B (MGX Business Continuity Solutions Pty Ltd) not relevant to specific survey – not part of the sample.

Enterprise New's (MGX Business Continuity Solutions Pty Ltd) capture the information received from the respondent but use the weight of enterprise A (the smaller enterprise's weight).

If the take over happened in quarter three, then quarter four's information would be collected and a questionnaire would be dispatched in quarter four to the new enterprise (MGX Business Continuity Solutions Pty Ltd). When the new sample is drawn then hopefully the take over would be recorded as it is observed in the real life situation and we would sample the newly created enterprise.

If in quarter four there was a response then the information for all variables would be captured for the survey for the new enterprise (MGX Business Continuity Solutions Pty Ltd) if not then the unit will be treated as non-response e.g.

Employment: Capture information received namely, **102 employees.**

Salaries and wages: Capture information received namely, **R 360 000 per quarter.**

Use weight of enterprise A.

Note – If the new USI code is 10 and the name has changed, the following procedure should be applied:

Changes to name, address and other contact details are to be changed on the SMS. Before any name and/or address changes are done the collection clerk should phone the respondent and enquire if this name and/or address change did not occur because of a demographic event. A demographic event occurs if the enterprise was sold, merged, taken over, split off and break up i.e. USI code 10. If no demographic event occurred then the collection clerk must allocate the applicable USI code according to the new information received from the respondent. Note that all demographic changes occurring in the sample should be confirmed with Systems of Registers before changes are made. During the next round of surveying the questionnaire must be posted with the new name and/or address. It is important to remember that a USI code cannot be allocated for a change in name and/or address.

The action required is to:

1. Inform the respondent that a completed questionnaire is required for enterprise A (CSX) if it is possible for the respondent to provide this information. Also, inform the respondent that no questionnaire for CSX would be completed in future;
2. Capture all relevant information required for future dispatch; and

3. Assign a USI code of 10 to the unit but if it is a non-response then leave it as a USI code of 00.

The above example is an example of a **structural change**. The key points are:

1. The selected unit permanently ceased generating income and incurring expenses during the reference period; and
2. The business activity of the selected unit is being carried out wholly by another business.

Identification and definition

These procedures apply to a selected unit when:

1. The business corresponding to the selected unit's identification name has ceased operating as a single entity; **and**
2. Some or all of the operations of the business were transferred, sold, taken over and/or merged with the operations of a real world business, which is in the scope of the survey.

Some examples include:

1. A restructuring of the business resulting in a merger of two or more business units within the enterprise group; and
2. A company selling its business and ceasing to earn income and/or incur expenses.

Example 1 describes a business restructure where the ownership does not change but the unit's availability for selection has changed. Example 2 is a common change of ownership.

A change of legal entity name is often an indicator that a transfer, sale, and/or merger have taken place.

In order to satisfy the structural change procedures criteria it is necessary to confirm with the respondent(s) that:

1. The selected unit permanently ceased generating income and incurring expenses before or during the reference period; **and**
2. The business activity of the selected unit is being carried out wholly, or in part, by another business/other businesses (identification name and contact details of the other business (es) are required).

For units that operated during the reference period sufficient information must be collected from the respondent to determine whether or not the characteristics of the business satisfy the definition of scope.

Data Processing

The same process will be applicable as for a take-over, merger, split off and break up. One enterprise will have a zero contribution to the estimates and no questionnaire will be dispatched to the respondent (if both units are in the sample). If only one unit is in the sample thus, if only enterprise A (the smaller enterprise) is included in the sample but enterprise B (the larger enterprise) is not included in the sample and the new enterprise name has changed as well as a change in its activity then enterprise A should be imputed for the remainder of the survey cycle. In a case where an enterprise takes over another enterprise, which is not in the sample and the effect is small then the results could be accepted. The effect of small could be described to be less than 5% within a subgroup and 1% within a group. This would usually be the case if the larger enterprise (enterprise B) were in the sample and the take-over happened with the smaller enterprise whereby enterprise B has taken over enterprise A.

SMS

Ensure the take over has the correct real world name and address information. Other details required for future dispatch should be recorded on the SMS.

USI codes

Corresponding data collection procedure	USI code	USI Description
Sampled unit has been sold to, merged with, taken over by, split off or broken up – Evident documentation available.	10	Unit indicates structural changes (e.g. take over, sold, merger etc.)

Estimation and future cycles

Procedure	Description	Contribution to survey estimates	Further dispatch action	
Sampled unit has been sold to, merged with, taken over by, split off or broken up – Evident documentation available.	Both enterprises are included in the sample.	<i>Enterprise A:</i> Capture as zero's and <i>Enterprise B:</i> Σ (Enterprise A + Enterprise B)	<i>Enterprise A:</i> No dispatch and <i>Enterprise B:</i> Dispatch and capture the relevant information (e.g. enterprise name, address, contact information etc.)	
	Only one enterprise is included in the sample. Enterprise A is included in the sample and enterprise B is not included in the sample)	Large effect on the estimates	<i>Enterprise A:</i> Impute	<i>Enterprise A:</i> No dispatch
		Small effect on the estimates	<i>New Enterprise:</i> Value	<i>New Enterprise:</i> Dispatch and capture the relevant information (e.g. enterprise name, address, contact information etc.)

Merger

Handle the same way as a take-over. The treatment of units as a result of an enterprise that is merged, is quite similar to that described under the heading take over with the exception of take-over being replaced by merger.

Sampled unit has been sold

Handle the same way as a take-over since the treatment of units as a result of an enterprise that is sold, is quite similar to that described under the heading take-over (Only one enterprise is included in the sample. Enterprise A is included in the sample and enterprise B is not included in the sample).

Break up or split off

Handle the same as take over The treatment of units as a result of an enterprise that breaks up or split off, is quite similar to that described under the heading take-over with the exception of take over been replaced by break up or split off.

(Only one enterprise is included in the sample. Enterprise A is included in the sample while enterprise B is not included in the sample).

(2) CLOSED DOWN/LIQUIDATED DURING THE SURVEY PERIOD – EVIDENT DOCUMENTATION AVAILABLE (USI 11)

When an enterprise is closed down or liquidated during the survey period, then the following can materialise:

Liquidated

Enterprises in liquidation are usually closed down although in some cases the final liquidation can take years. As the activities of such enterprises are normally suspended (they are not productive), these enterprises are recommended to still form part of the sample and are imputed if the data cannot be collected from the specific enterprise, till the final liquidation of the enterprise.

If the selected unit is in liquidation, treat the **date** it went into liquidation as the date it ceased conducting business. It is expected that this unit will not continue to operate in the near future.

- In the case of liquidation the date of liquidation should be confirmed with the respondent and if possible collect the data during the liquidation process, while if the enterprise has merged and its operations continue we have to collect the data on that enterprise. The main reason why a change was observed in the merged unit is not primarily because of the liquidation of the enterprise but because of the merger that took place. The sample maintenance procedure that should be followed in such a case is the one on mergers.
- If the enterprise did not merge but the data cannot be collected then imputation can be done only when the liquidation falls within the reference period. It should be noted that if the unit went into liquidation during the reference period and it is not possible to obtain data it should be imputed.

Example

A statistical unit selected in a certain survey has the following sample attributes:

Frame identification name (Legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638
Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Classification: SIC 8899 (Other business activities n.e.c.).

Turnover: R 61 039 185,43131.

Mr. Phillips (the accountant of MGX Business Continuity Solutions Pty Ltd, phoned to say that he had received a survey questionnaire but he didn't believe it was relevant, as the company had closed down their business.

The processing clerk asked if the name and address information were correct. The respondent confirmed that they were. The processing clerk then asked when MGX Business Continuity Solutions Pty Ltd ceased generating income and incurring expenses. Mr. Phillips indicated it was mid July 2002. The clerk asked whether the MGX Business Continuity Solutions Pty Ltd had ceased operating because its operations had been sold to another business. Mr. Phillips said this was not the case. The clerk asked if there was any intention for MGX Business Continuity Solutions Pty Ltd to recommence business in the near future. Mr. Phillips said that the company is not going to bother deregistering the company name but the company didn't expect to go back into business again. This confirmed to the processing clerk that the unit satisfied the closed down/liquidated criteria as it had not generated income or incurred expenses during the reference period and there was no intention to recommence operations.

The clerk then asked if the business activity had been sold to another entity. Mr. Phillips replied that it had not. The clerk then requested a written letter, fax or e-mail confirming the conversation.

Application:

Impute the information for the enterprise (MGX Business Continuity Solutions Pty Ltd).

Respondent contact

In order to satisfy the closed down/liquidated criteria it is necessary to confirm with the respondent that:

1. The label name corresponds to what was the correct name of the selected unit;
and
2. There is no intention for the selected unit to operate in the future.
3. The selected unit ceased generating income and incurring expenses during the reference period. Note that it is necessary to request written confirmation from the respondent.

For units that operated during the reference period, sufficient information must be collected from the respondent to determine whether or not the characteristics of the business satisfy the definition of scope.

Data processing

If the necessary documentation is received from the respondent and sufficient information is available then the respondent is not required to complete the survey questionnaire.

The respondent is only required to complete the survey questionnaire for the reference period, if there is insufficient information on the unit, but if a letter of confirmation is received this unit can be treated as a closed down/liquidated procedure during the survey period where evident information is available. These units will be imputed.

If the unit ceased operating during the reference period and the date when the selected unit ceased operations is known but no data is forthcoming from the respondent the unit can be imputed using this date for the portion of the reference period in which it operated.

SMS

The closed down/liquidation should be captured on the Survey Management System so that it can be used in the estimation and imputation of the survey. This will make it possible for all survey areas to have access to this information thereby ensuring that duplication of work does not happen. Thus, this could serve in the assistance of running the survey and keeping survey areas informed about all other information that survey areas have received from respondents.

USI codes

Corresponding data collection procedure	USI code	USI Description
Closed down/Liquidated during the survey period – Evident documentation available	11	Unit indicates ceased operations (e.g. closed down /liquidation)

Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Closed down/Liquidated during the survey period – Evident documentation available	Impute	No dispatch

Note: As a general rule, survey areas should impute only once if the sample is annual and not more than twelve times if the sample is monthly, etc. but the only way to address the problem is by drawing samples more frequently in future.

This is a clear example of a closed down/liquidated unit. The key points are:

1. The respondent was talking to the processing clerk about the selected unit (identification) and name of business was correct;
2. The business had no income and no expenses during the reference period;
3. The operations of the business had not been merged with, transferred to or sold to another unit; and
4. The business was not likely to operate in the future.

The action required is to:

1. Inform the respondent the questionnaire doesn't need to be completed;
2. Capture all relevant information on the SMS; and
3. Assign a USI code of 11 if written proof is received and sufficient unit information is collected.

[Closed down](#)

Handle the same way as liquidated. The treatment of units as a result of an enterprise that closed down during the survey period, is quite similar to those described under the heading liquidated with the exception of liquidated being replaced by closed down.

(3) FINALLY LIQUIDATED – EVIDENT DOCUMENTATION AVAILABLE (BEFORE THE SAMPLE WAS DRAWN) (USI 09)

When an enterprise is finally liquidated before the sample was drawn, then the following can materialise:

Enterprises that are finally liquidated and evident documentation is available and if this liquidation happened before the sample was drawn – thus the activities of such enterprises were suspended (they are not productive) a while back – it is recommended that the weights be adjusted to compensate for this type of occurrence (non-response). This procedure will prevent numerous enquiries and investigations and in this case the questionnaires would not be dispatched again to the specific enterprise. If the liquidation falls outside the reference period, it should be treated as non-response.

Example

A statistical unit selected in a certain survey has the following sample attributes:

Frame identification name (Legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638
Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Classification: SIC 8899 (Other business activities n.e.c.).

Turnover: R 61 039 185,43131.

Mr. Phillips (the accountant of MGX Business Continuity Solutions Pty Ltd, phoned to say that he had received a survey questionnaire but he didn't believe it was relevant, as the company had closed down their business almost a year ago.

The processing clerk asked if the name and address information were correct. The respondent confirmed that it was. The processing clerk then asked when MGX Business Continuity Solutions Pty Ltd ceased generating income and incurring expenses. Mr. Phillips indicated it was mid-August 2001. The clerk asked whether the MGX Business Continuity Solutions Pty Ltd had ceased operating because its operations had been sold to another business. Mr. Phillips said this was not the case. The clerk asked if there was any intention for MGX Business Continuity Solutions Pty Ltd to recommence business in the near future. Mr. Phillips said that the company is not going to bother deregistering the

company name but the company didn't expect to go back into business again. This confirmed to the processing clerk that the unit satisfied the finally liquidated/closed down criteria as it had not generated income or incurred expenses before the reference period and there is no intention to recommence operations. Thus MGX Business Continuity Solutions Pty Ltd was liquidated before the sample was drawn and this change didn't appear on the sampling frame of June 2002.

The clerk then asked if the business activity had been sold to another entity. Mr. Phillips replied that it had not. The clerk then requested a written letter, fax or e-mail confirming the conversation.

Application:

Adjust the weights for non-response. (See page 26)

Respondent contact

In order to satisfy the finally liquidated procedures it is necessary to confirm with the respondent that:

1. The label name corresponds to what was the correct name of the selected unit;
and
2. There is no intention for the selected unit to operate in the future.
3. The selected unit ceased generating income and incurring expenses before the reference period. Note that it is necessary to request written confirmation of this from the respondent.

For all units that ceased operation before the sample was drawn (before the reference period), sufficient information must be collected from the respondent to capture this information and the characteristics of the business on the SMS.

Data processing

If the necessary documentation is received from the respondent and sufficient information is available then the respondent is not required to complete the survey questionnaire and the unit could be seen as finally liquidated. The weights would be adjusted for these units.

SMS

The finally liquidated enterprise should be captured on the Survey Management System so that it can be used in the estimation and imputation of the survey. This will make it possible for all survey areas to have access to this information and ensuring that duplication of work does not happen. Thus this could serve in the assistance of running the survey and keeping survey areas informed about other information that survey areas have received from respondents.

USI codes

Corresponding data collection procedure	USI code	USI Description
Finally liquidated – Evident documentation available (before the sample was drawn)	09	Unit indicates ceased operations thus finally liquidated

Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Finally liquidated – Evident documentation available (before the sample was drawn)	Adjust weights for non-response	No dispatch

This is a clear example of a finally liquidated enterprise with evident information available. The key points are:

1. The respondent was talking to the processing clerk about the selected unit (identification) and name of business and correct;
2. The business had no income and no expenses before the sample was drawn;
3. The operations of the business had not been merged with, transferred to or sold to another unit; and
4. The business was not likely to operate in the future.

The action required is to:

1. Inform the respondent the questionnaire should not be completed;
2. Capture all relevant information on the SMS; and
3. Assign a USI code of 09, if written proof is received and sufficient unit information is collected.

(4) CLOSED DOWN – EVIDENT DOCUMENTATION AVAILABLE (BEFORE THE SAMPLE WAS DRAWN) (USI 07)

Handle the same way as finally liquidated – evident documentation available (before the sample was drawn) (USI 09). The treatment of units as a result of an enterprise that closed down before the sample was drawn, is quite similar to those described under the heading dealing with finally liquidated with the exception that finally liquidated be replaced by closed down.

(5) DORMANT – EVIDENT DOCUMENTATION AVAILABLE (USI 08)

Handle the same way as liquidated. The treatment of units as a result of an enterprise that is dormant during the survey period, is quite similar to those described under the heading dealing with liquidated with the exception that liquidated be replaced by dormant.

(6) UNDELIVERED QUESTIONNAIRE (USI 01)

Survey areas should attempt to collect untraceable enterprises, especially if these enterprises are the size of group 1 enterprises. In the absence of any knowledge about these enterprises the assumption is made that the enterprise is still alive and relevant information of the unit should be collected for the relevant period.

If the questionnaire was returned to Stats SA, marked as ‘undelivered mail’ the processing clerk should start to follow the necessary ‘trace’ procedures ([see also chapter 4](#)) in the process to find the correct information about this enterprise.

Application:

The following steps should be taken by subject matter areas with respect to undelivered mail (untraceable names and addresses) – tracing procedures:

1. Determine the **enterprise number** and use this in an attempt to find a usable telephone number in the Business Sampling Frame or at SARS.
2. Try to find **another name** (e.g. **trading name**), which is related to the enterprise being sought.
3. If there is **more than one trading name**, compare with **possible trading names** on the **SARS database**. Contact information of the branch with the **largest turnover** should be determined.
4. According to SARS, a specific branch may be **dormant**. Information on the **status** of this branch can then be obtained from one of the **non-dormant** branches.
5. **Consult telephone books** and other directories such as **Brabies**.
6. Request **Telkom’s** assistance at the number **1023**. Sometimes their guides do contain the name in **abbreviated form**.
7. Sometimes **Telkom cannot trace** the exact name but can supply **similar names**, which may **lead to identification**.
8. Contact the **auditor (accountant)** of the enterprise if his/her information is available.
9. **Verify the information** about the unit through the use of the **SMS**. It might be possible that another survey area was fortunate enough to get the applicable contact information etc.
10. Try any other method available.

It should be noted that although undelivered mail does not represent a structural change within an enterprise, it has a serious effect on sample maintenance and the timeliness of

sample results. If an enterprise however, closes down or merges with another enterprise, undelivered mail can be an indication of a structural change. It is therefore very important that all possible measures and actions be taken to clarify the position of each enterprise involved in the problem of undelivered mail.

If no information could be obtained about the specific enterprise then the enterprise should be treated as untraceable and a USI 04 code should be assigned.

Data processing

If the necessary tracing procedures are followed and the correct information for the enterprise could be obtained then the questionnaire would be dispatched but if no other information is available for this specific unit then the questionnaire would not be dispatched again in the following survey period. It should however be noted that the survey area could decide to re-dispatch the questionnaires of those enterprises previously marked as undelivered.

If no information of the unit could be collected then the enterprise would be treated as untraceable.

SMS

The undelivered questionnaires should be captured on the Survey Management System so that the information can be used in the estimation and imputation of the survey also it will make it possible for all survey areas to have access to this information. Thus this could serve in the assistance of running the survey and keeping survey areas informed about other information that survey areas have received from respondents.

USI codes

Corresponding data collection procedure	USI code	USI Description
Undelivered questionnaire	01	If minimum 'trace' procedures was followed then treat as untraceable

Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Undelivered questionnaire (If no information received)	Impute	* No dispatch

*** Note:** surveys can choose to re-dispatch to all of the units that fall in the category of undelivered questionnaire.

Note: As a general rule survey areas should impute only once if the sample is annual and not more than twelve times if the sample is monthly, etc. but the only way to address the problem is by drawing samples more frequently in future.

If all the steps are followed and no information is received then this is a clear example of an undelivered questionnaire. The key point is:

1. The processing clerk followed all the tracing procedures and no additional information could be obtained;

The action required is to:

1. Capture all relevant information on the SMS.
2. Assign a USI code of 01 if the necessary information is still outstanding; and
3. Assign a USI code of 04 if the necessary tracing procedures were followed.

(7) UNTRACEABLE – PROCESS OF INVESTIGATION FINALISED (USI 04)

Survey areas should attempt to collect untraceable enterprises, especially if these enterprises are the size of group 1 enterprises. In the absence of any knowledge about these enterprises the assumption is made that the enterprise is still alive and the unit should be imputed for the relevant period.

Example

A statistical unit selected in a certain survey has the following sample attributes:

Frame identification name (legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638

Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Classification: SIC 8899 (Other business activities n.e.c.).

Turnover: R 61 039 185,43131.

The questionnaire was returned to Stats SA, marked as ‘undelivered mail’. The processing clerk started to follow the necessary ‘trace’ procedures ([see also chapter 4](#)) in the process to find the correct information about this enterprise.

Application:

The subject matter areas should follow the minimum tracing procedures with respect to undelivered mail (untraceable names and addresses) – as set out below:

- 1 Determine the **enterprise number** and use this in an attempt to find a usable telephone number in the Business Sampling Frame or at SARS.
- 2 Try to find **another name** (e.g. **trading name**), which is related to the enterprise being sought.
- 3 If there is **more than one trading name**, compare with **possible trading names** on the **SARS database**. Contact information of the branch with the **largest turnover** should be determined.
- 4 According to SARS, a specific branch may be **dormant**. Information on the **status** of this branch can then be obtained from one of the **non-dormant** branches.
- 5 **Consult telephone books** and other directories such as **Brabies**.
- 6 Request **Telkom's** assistance at the number **1023**. Sometimes their guides do contain the name in **abbreviated form**.
- 7 Sometimes **Telkom cannot trace** the exact name but can supply **similar names**, which may **lead to identification**.
- 8 Contact the **auditor (accountant)** of the enterprise if his/her information is available.
- 9 **Verify the information** about the unit through the use of the **SMS**. It might be possible that another survey area was fortunate enough to get the applicable contact information etc.
- 10 Try any other method available.

It should be noted that although undelivered mail does not represent a structural change within an enterprise, it has a serious effect on sample maintenance and the timeliness of sample results. If an enterprise however, closes down or merges with another enterprise, undelivered mail can be an indication of a structural change. It is therefore, very important that all possible measures and actions should be taken to clarify the position of each enterprise involved in the problem of undelivered mail.

If no information could be obtained about the specific enterprise then impute the information for the enterprise (MGX Business Continuity Solutions Pty Ltd).

Data processing

Since no other information is available for this specific unit then the questionnaire would not be dispatched again in the following survey period. These enterprises would be seen as untraceable, if all the necessary tracing procedures were followed.

These units will be imputed.

SMS

The untraceable units should be captured on the Survey Management System so that it can be used in the estimation and imputation of the survey also it will make it possible for all survey areas to have access to this information. Thus this could serve in the assistance of running the survey and keeping survey areas informed about other information that survey areas have received from respondents.

USI codes

Corresponding data collection procedure	USI code	USI Description
Untraceable	04	Minimum 'trace' procedures were followed and no additional information could be collected.

Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Undelivered questionnaire	Impute	No dispatch

Note: As a general rule survey areas should impute only once if the sample is annual and not more than twelve times if the sample is monthly, etc. but the only way to address the problem is by drawing samples more frequently in future.

If all the steps are followed and no information is received then this is a clear example of an undelivered questionnaire. The key point is:

1. The processing clerk followed all the tracing procedures and no additional information could be obtained;

The action required is to:

1. Capture all relevant information on the SMS; and
2. Assign a USI code of 04 if the necessary tracing procedures were followed.

(8) SELECTED UNIT SITUATED OUTSIDE THE BORDERS OF SOUTH AFRICA (USI 16)

Handle the same way as untraceable. The treatment of units as a result of an enterprise that is situated outside the borders of South Africa, is quite similar to those described under the heading untraceable with the exception of untraceable been replacing selected unit situated outside the borders of South Africa.

(9) DUPLICATE – DIFFERENT SAMPLED UNITS, BUT SAME BUSINESS (USI 28)

Duplicate procedures apply when two or more sampling units are found to correspond to the same real world business and only one real world business ever existed.

Duplicates are usually identified when a respondent receives more than one questionnaire addressed to the same real world business. Further investigation may identify additional duplicates on the frame, which have not been selected in the sample of the specific survey. Similarly a quality check of the frame prior to dispatch may identify a number of duplicates. Irrespective of how duplicates are identified the following definitions will apply.

The **primary duplicate** is the sampling unit chosen to represent the real world business in a survey. The primary duplicate is the sampling unit with the largest measure of size on the sampling frame thus the highest turnover on the frame. If the operations of the real world business have changed then the appropriate procedure should apply (e.g. if the real world unit ceased operations then the appropriate procedures should be applied, if the real world unit is liquidated then the appropriate liquidation procedures should be applied, etc.)

All other sampling units corresponding to the real world business will be referred to as **secondary duplicates**. If the duplicated units are classified in the same industry group (SIC), one enterprise must be treated as zero response (secondary duplicate). If the duplicate unit has a different classification, use correct classification of the enterprise (primary duplicate) and treat the relevant enterprise as zero response (secondary duplicate) thus capture as zeros.

Related and special cases

Transfer, sale, takeover and/or merger has taken place:

If the business corresponding to the selected unit's identification name has closed down or the operations have been continued by another entity (a takeover), the unit should be treated according to the outlined procedures.

Respondent contact

In order to satisfy duplication criteria it is necessary to confirm with the respondent that:

1. At least one label name corresponds (minor enhancements are allowed) to the correct real world name of the selected unit;
2. The real world business has not been involved in the transfer, sale and/or merger of business operations which would fully explain multiple units on the frame; **and**
3. The real world business continues to operate.

Sufficient information must be collected from the respondent to determine whether or not the characteristics of the business satisfy the definition of scope.

Data processing

Primary duplicate:

Collect data from the real world business and process under the unique identifier (e.g. EN) of the primary duplicate. It is important to remember to collect data only from a primary duplicate.

Secondary duplicate:

Do not process data under the unique identifier (e.g. EN) of the secondary duplicate. The sampling unit, which has the largest measure of size, thus the primary duplicate in the sample, will be used as input to the survey's data. Data should only be collected for the primary duplicate in future survey cycles. For the secondary duplicate no questionnaire would be re-dispatched in future and should be captured as zeros. The reason for this is to avoid over estimation.

SMS

Ensure the primary duplicate has the correct real world name and address information and that the unique identifier of any secondary duplicates is recorded. Other details required for future dispatch should be recorded on the SMS.

USI codes

Corresponding data collection procedure	Description	USI code	USI Description
Duplicate – different sampled units, but same business	Primary Duplicate	22	Complete information received, checked, compared and captured. Received completed questionnaire from respondent (USI 22)
	Secondary Duplicate	28	Duplicate (when two or more selected units correspond to the same real world business and only one real world business ever existed)

Estimation and future cycles

Procedure	Description	Contribution to survey estimates	Further dispatch action
Duplicate – different sampled units, but same business	Primary duplicate	Value	Dispatch
	Secondary duplicate	Zero	No Dispatch

Example

A statistical unit selected in a certain survey has the following sample attributes:

Frame identification name (Legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638
Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Sampling unit unique identifier: EN 15356983.

Classification: SIC 8899 (Other business activities n.e.c.).

Turnover: R 61 039 185,43131.

Mr. Phillips phoned because he had received two questionnaires for the same business and wanted to inform Stats SA of the duplication.

The processing clerk asked if the identification names on both questionnaires were exactly the same. Mr. Phillips said that one questionnaire was perfectly correct (MGX Business Continuity Solutions Pty Ltd) and the other was correct except it spelled out the MGX in full rather than using the abbreviation.

The processing clerk checked that there have not been any demographic events/structural changes during the reference period, that the other label information was correct and that the unit was in scope of the Quarterly Financial Statistics (QFS) survey.

This is an example of when to apply the duplicate – different sampled units, but same business procedures. The key points are:

1. The label name corresponded with the real world name of the selected unit;
2. The real world business has not been involved in a transfer, sale and/or merger of business operations;

3. The real world business continues to operate; and
The real world business' activities are within the scope of the survey.

The action required is to:

1. Inform the respondent to complete the questionnaire with the correct name (EN 12136123) and request that both questionnaires be returned in the same envelope;
2. Assign a USI code of 28 to the unit with the enterprise number: EN 15356983, record the EN of the secondary duplicate on the SMS, and ensure the name and address information is correct; and
3. Assign a USI code of 22 to the unit with the enterprise number EN 12136123 and record the EN of the primary unit on SMS, this will be captured as an actual response.

(10) INFORMATION OF ONE SAMPLED UNIT INCLUDED IN THE INFORMATION OF ANOTHER SAMPLED UNIT (WITHIN THE SAMPLE) (USI 12)

Handle the same way as duplicate – different sampled units, but same business. The treatment of units as a result of an enterprise information (sampled unit) included in the information of another sampled unit, within the sample, is quite similar to those described under the heading duplicate – different sampled units, but same business except that duplicate – different sampled units, but same business will be replaced by information of one sampled unit included in the information of another sampled unit (within the sample).

(11) INFORMATION OF ONE SAMPLED UNIT INCLUDED IN THE INFORMATION OF ANOTHER UNIT (OUTSIDE THE SAMPLE) (USI 13)

These procedures apply when the business operations in the real world are different from the unit attributes in the sample and the differences result in the selected unit becoming out of scope of the survey. Out of scope procedures is not applicable when the unit is incorrectly classified. Domain analysis will be used for incorrectly classified units within the sample and where the unit is incorrectly classified and the re-classified enterprise falls outside the sample scope, then the specific unit will be treated according to the procedures outlined in this manual. Each survey area in consultation with the relevant Survey Statistician determines the survey scope.

Selected units no longer in scope of the survey will be referred to as **out of scope** thus outside the sample. If a real world business is out of scope of the survey, the selected unit should be treated as **out of scope** regardless of whether other differences are found. For example, if the real world business is a seasonal operator but it is no longer in scope of the survey and the units information is included in another unit's information the selected unit should be treated under information of one sampled unit included in the information of another unit (outside the sample) procedures.

Respondent contact

In order to satisfy the information of one sampled unit included in the information of another unit (outside the sample) procedures criteria, it is necessary to confirm with the respondent that:

1. The label name corresponds (minor enhancements are allowed) to the correct real world name of the selected unit; **and**
2. The selected unit (as identified by the identification name) continues to operate in the real world and all the relevant information on the other enterprise (unit) should be collected.

Sufficient information has to be collected from the respondent to determine whether or not the characteristics of the business satisfy the definition of scope.

Data processing

Included in the information of another unit (outside the sample):

The respondent is not required to fill in the survey questionnaire. If other changes to the business operations apply then collect necessary data to capture on the SMS and for future dispatch. The weights would be adjusted in this case.

SMS

Capture name and/or address enhancements necessary for Business Sampling Frame feedback or further dispatch. Record the reason why the selected unit is out of scope of the survey and any associated information. For example, if the selected unit is found to be in an industry different from that recorded in the sample either the correct activity description or SIC code for the business should be captured and if the activity is within the scope of the survey then domain analysis will be applied.

USI codes

Corresponding data collection procedure	USI code	USI Description
Information of one sampled unit included in the information of another unit (outside the sample)	13	Unit unable to respond because the unit's information is included in another unit that falls outside the sample

Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Information of one sampled unit included in the	Adjust weights for non-response	No dispatch

information of another unit (outside the sample)		
---	--	--

Application:

Adjust the weights for non-response.

[See finally liquidated](#) – Evident documentation available (before the sample was drawn) for description of the weighting procedure.

Example

A statistical unit selected in the manufacturing survey has the following sample attributes:

Frame identification name (Legal name): MGX Business Continuity Solutions Pty Ltd.

Trading as: MGX Business Continuity Solutions Pty Ltd.

Mailing address: P.O. Box 2638
Midrand
South Africa
2018

Sampling unit unique identifier: EN 12136123.

Classification: SIC 3590 (Manufacture of office, accounting and computing machinery).

Turnover: R 61 039 185,43131.

Mr. Phillips, the accountant of MGX Business Continuity Solutions Pty Ltd, called to say that the company would not be able to complete the survey questionnaire received from Stats SA because the financial statements etc. of this enterprise is compiled and reported with another business unit namely MGX Pty Ltd.

The processing clerk confirmed the identification name of the business was correct and that there had been no change in business operations during the reference period and he/she collected information on the out of scope unit (MGX Pty Ltd).

The key points, that identify this an occurrence of Information of one sampled unit included in the information of another unit (outside the sample) example are:

1. The label name corresponds to the real world name of the selected unit,
2. The selected unit's activity is out of scope for the survey, SIC 8899 – Other business activities n.e.c.

The action required is to:

1. Inform the respondent the questionnaire does not need to be completed; and
2. Assign a USI code of 13 to the unit.

(12) FOREIGN BUSINESSES, E.G. BRITISH AIRWAYS, AIR FRANCE. THE FINANCIAL INFORMATION IS NOT AVAILABLE FOR THE SOUTH AFRICAN BUSINESS (USI 15)

Handle the way same as untraceable. The treatment of units as a result of an enterprise, foreign businesses, e.g. British Airways, Air France whose financial information is not available for the South African business, is quite similar to that described under the heading untraceable with the exception that untraceable will be replaced by foreign businesses, e.g. British Airways, Air France whose financial information is not available for the South African business. No questionnaire should be dispatched and the unit should be imputed.

(13) RECEIVED COMPLETED QUESTIONNAIRE FROM RESPONDENT BUT RE-CLASSIFIED – WITHIN THE SAMPLE - DOMAIN VARIABLE (USI 17)

It could be assumed that misclassifications exist on the part of both parties and are of equal size and so they cancel each other out. The assumption should of course be presented openly in a way that the user of the statistics understands.

Methods:

1. Count these enterprises in the activity from which they were sampled. This will of course yield a classification error, but no enterprise information is lost.
2. Alternatively the sample activity is kept for all enterprises except those for which the misclassification is detected. For the latter, the correct activity is assigned. For the estimation the weight $\frac{N_h}{n_h}$ used is that of the sampling stratum. In this case the classification error becomes slightly less but the variance will be larger because of the correctly classified enterprises (if they are sampled). No enterprise information is lost if the survey is economy-wide. Domain analysis.

Sector oriented surveys:

These types of surveys differ from that of the economy-wide ones in that the information collected is sector specific. This means that seriously misclassified enterprises most likely become **non-responses**. Suppose for instance that the misclassification is less serious and that “Hotel & Restaurants” are surveyed and that an enterprise is sampled from “Hotels” but the correct classification is “Restaurants” (*a misclassification on 3-digit level*) and the presentation of the statistics is on a higher SIC- level then this type of misclassification would not have such a large effect on the classification error.

Economy-wide surveys:

1. Use the incorrectly classified enterprises as correctly classified and then handle them like all the other enterprises in the industry;

or

2. Alternatively, reclassify for size group 1 – Reclassify the unit to the correct industry group (SIC) but use the weight of the original sampled enterprise thus introduce a **domain variable**, which will be used in **domain analysis**. This will mean using information as received by the original sampled enterprise (the respondent) for Size group 1 the units (enterprise), with a drawing probability of 1 and **capture** the information in the **correct classification** thus no information will be lost and the sampled weight should be used. In the case of size group 1 the wrongly classified enterprise should be changed to the correct classification. In the case of size groups 2, 3 and 4 the **wrong classification** should be **kept as is**. This is applicable also when the classification changes from one reference period to the next reference period. For **size group 2, 3, 4** an acceptable **imputation** method should be used or alternatively treat it as **non-response (adjust weight)**.

Note: All survey areas should include a question on the main activity of the enterprise that is surveyed because this will assist in determining the misclassifications in the sample.

SMS

Received completed questionnaire from respondent but re-classified – within the sample - domain variable should be captured on the database so that it can be used in the estimation and imputation of the survey. Also, misclassifications should be noted on the SMS because it will make it possible for all survey areas to have access to this information. This could serve in the assistance of running the survey and keeping survey areas informed about other information that survey areas have received from respondents.

USI codes

Corresponding data collection procedure	USI code	USI Description
Received completed questionnaire from respondent but re-classified – Within the sample - domain variable (USI 17)	17	Misclassified unit

Estimation and future cycles

Procedure	Description	Contribution to survey estimates	Further dispatch action
Received completed questionnaire from respondent but re-classified – Within the sample - domain variable (USI 17)	Economy-wide surveys	Value	Dispatch
	Sector oriented surveys	Value	Dispatch

(14) RECEIVED COMPLETED QUESTIONNAIRE FROM RESPONDENT BUT RE-CLASSIFIED – OUTSIDE THE SAMPLE (USI 18)

Sector oriented surveys:

Application:

Adjust the weights for non-response.

[See finally liquidated](#) – Evident documentation available (before the sample was drawn) for the description of the weighting procedure.

Economy-wide surveys:

Application:

Economy-wide surveys would in almost all the cases be treated under the domain analysis as outlined under received completed questionnaire from respondent but re-classified – within the sample - domain variable (USI 17), but cases that still fall outside the sample could be dealt with either with the appropriate imputation techniques or they could be treated as non-response (Adjust weights for non-response).

SMS

The received completed questionnaire from respondent but re-classified – outside the sample should be captured on the Survey Management System so that it can be used in the estimation and imputation of the survey also it will make it possible for all survey areas to have access to this information. This could serve in the assistance of running the survey and keeping survey areas informed about other information that survey areas have received from respondents.

Note: All the necessary changes should be done on the SMS.

USI codes

Corresponding data collection procedure	USI code	USI Description
Received completed questionnaire from respondent but re-classified – Outside the sample	18	Misclassified unit

Estimation and future cycles

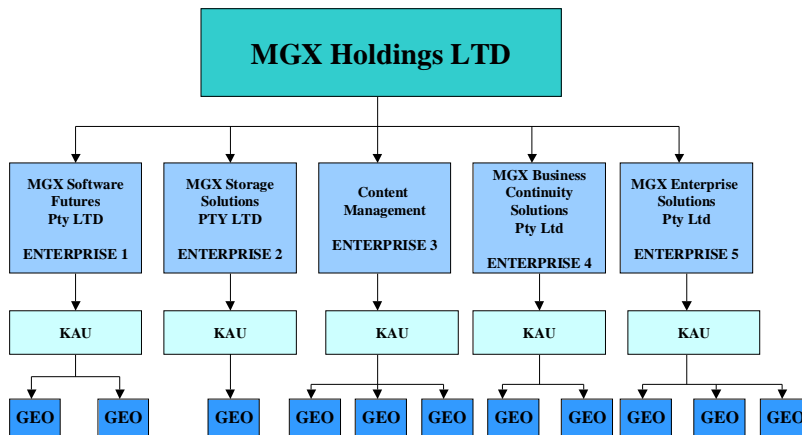
Procedure	Description	Contribution to survey estimates	Further dispatch action
Received completed questionnaire from respondent but re-classified – Outside the sample	Economy-wide surveys	Impute or Adjust weights for non-response	Dispatch
	Sector oriented surveys	Adjust weights for non-response	No dispatch

(15) SAMPLED UNIT UNABLE TO SUPPLY INFORMATION ON AN ENTERPRISE LEVEL (USI 05)

If there is a selected unit that is known to be ‘live’ but is unable to report on the structural level that is required, then an observation unit should be created on the SMS.

Examples

Example 1



The structure of MGX Holdings LTD can be seen above. This Holding Company consist of five enterprises, each consisting of a KAU structure and their relevant GEO structures. MGX Holdings LTD was not on the Business Sampling Frame (BSF) when the sampling frame was created. A sample was drawn and the following three Enterprises were included in the sample: Enterprise 1, Enterprise 2 and Enterprise 5.

Questionnaires should be dispatched to Enterprise 1, Enterprise 2 and Enterprise 5.

Application:

Two of the three units responded but the accountant of one of the units, MGX Enterprise Solutions Pty Ltd (Enterprise 5), phoned to say that the unit would not be able to provide information for that enterprise because they have a centralised payroll system and the financial information and statements of the enterprise would be included with the information of the enterprise group's information, comprising all five enterprises (MGX Holdings LTD).

Capturing:

Capture the information of the two units that responded.
Enterprise 5 (MGX Enterprise Solutions Pty Ltd) would be treated, as non-response and the information of this unit would be imputed for this period.

Note: It is not acceptable to use information on the enterprise groups e.g. Financial Statements of the group MGX Holdings LTD and treat all three enterprises as responded.

Example 2

The structure of MGX Holdings LTD can be seen above in example 1. This holding company consists of five enterprises, each consisting of a KAU structure and their relevant GEO structures. MGX Holdings LTD was not on the Business Sampling Frame (BSF) when the sampling frame was created. A sample was drawn and the following three Enterprises were included in the sample: Enterprise 1, Enterprise 2 and Enterprise 5.

Questionnaires should be dispatched to Enterprise 1, Enterprise 2 and Enterprise 5.

The accountant of Enterprise 5 (MGX Enterprise Solutions Pty Ltd) phoned to say that the company would not be able to report on the structural level but would be able to provide Stats SA with figures on a lower level – the production level. The processing clerk requested the information on the three units (observation units) with the necessary contact details and addresses to dispatch questionnaires to these units.

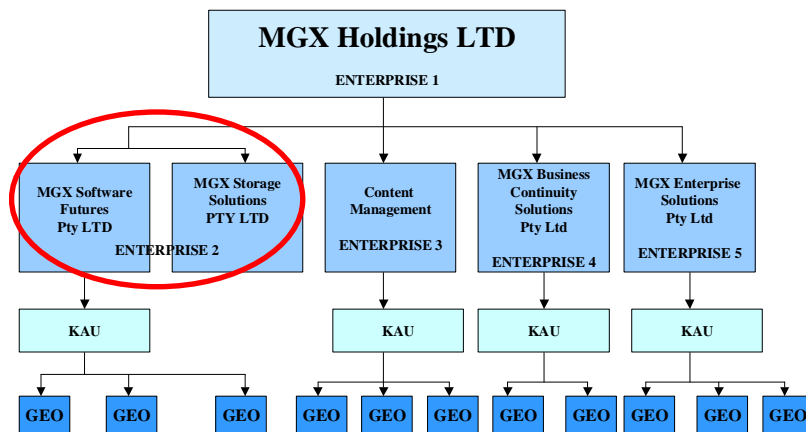
Capturing:

Capture the information of the two units that responded.

Create observation units for Enterprise 5 (MGX Enterprise Solutions Pty Ltd) on the SMS and make sure that all relevant information is captured on the SMS. The contact addresses of the observation units will be used in future to dispatch the questionnaires. If all three of these observation units responded then the summation of all three would be taken and captured as one enterprise, namely enterprise 5. Thus Enterprise 5 = Σ (observation unit ₁ + observation unit ₂ + observation unit ₃).

Example 3

In some cases enterprises (the holding companies) cannot supply the relevant information on enterprise level. Each unit of such a company (enterprise) can supply its own information. If the holding company is an element of the sample and cannot provide the necessary information the unit should be treated as non-response. Create an observation unit on the SMS for those enterprises that cannot provide information on the enterprise level.



The structure of MGX Holdings LTD can be seen above. This holding company consist of five enterprises, four enterprises consist of a KAU structure and their relevant GEO structures. MGX Holdings LTD was on the Business Sampling Frame (BSF) when the sampling frame was created. A sample was drawn and the following three Enterprises were included in the sample: Enterprise 1 (MGX Holdings LTD), Enterprise 2 (MGX Software solutions Pty Ltd and MGX Storage Solutions Pty Ltd) and Enterprise 5 (MGX Enterprise Solutions Pty Ltd). Note that Enterprise 2 (MGX Software solutions Pty Ltd and MGX Storage Solutions Pty Ltd) is grouped as one enterprise on the Business Sampling Frame.

Questionnaires should be dispatched to Enterprise 1, Enterprise 2 and Enterprise 5.

For Enterprise 1, the holding company, only the basic information is requested .The group figures should not be requested because then we will overestimate. If they can't provide Stats SA with information of the holding company only then should this unit be treated as a non-response and the unit should be imputed.

If the scenario is similar to the case of Enterprise 2, then create for both enterprises an observation unit and collect the information separately because in the real world this is two separate operations that each function on their own. It is important that the processing clerk request all the information of the two units (observation units) with the necessary contact details and addresses to dispatch questionnaires to these units.

The accountant of Enterprise 5 (MGX Enterprise Solutions Pty Ltd) phoned to say that the company would not be able to report on the structural level but would be able to provide Stats SA with figures on a lower level – the production level. The processing clerk requested all the information of the three units (observation units) with the necessary contact details and addresses to dispatch questionnaires to these units.

Capturing:

Enterprise should be imputed.

Enterprise 2 = Σ (observation unit ₁ + observation unit ₂).

1. Create observation units for **Enterprise 5** (MGX Enterprise Solutions Pty Ltd) on the SMS and make sure that all relevant information is captured on the SMS. The contact addresses of the observation units will be used in future to dispatch the questionnaires.
2. If all three of these observation units responded then the summation of all three would be taken and captured as one enterprise, namely Enterprise 5. Thus Enterprise 5 = Σ (observation unit ₁ + observation unit ₂ + observation unit ₃). If one of these observation units didn't respond then the unit will be treated as a non-response and the information for this unit will be imputed.

USI codes

Corresponding data collection procedure	USI code	USI Description
Sampled unit unable to supply information on an enterprise level	05	Unit cannot provide information on the structural level

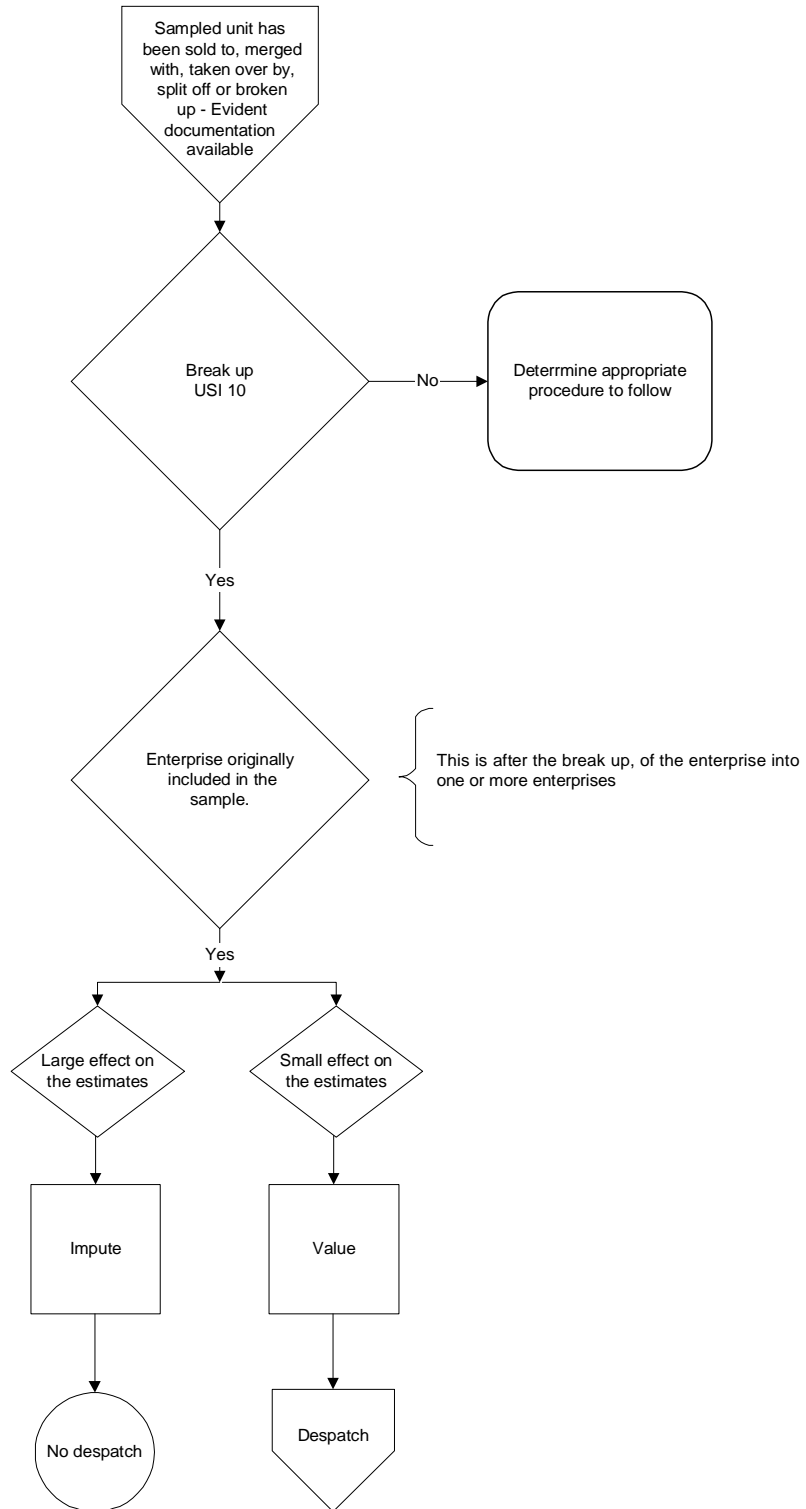
Estimation and future cycles

Procedure	Contribution to survey estimates	Further dispatch action
Sampled unit unable to supply information on an enterprise level	Value/Impute	Dispatch

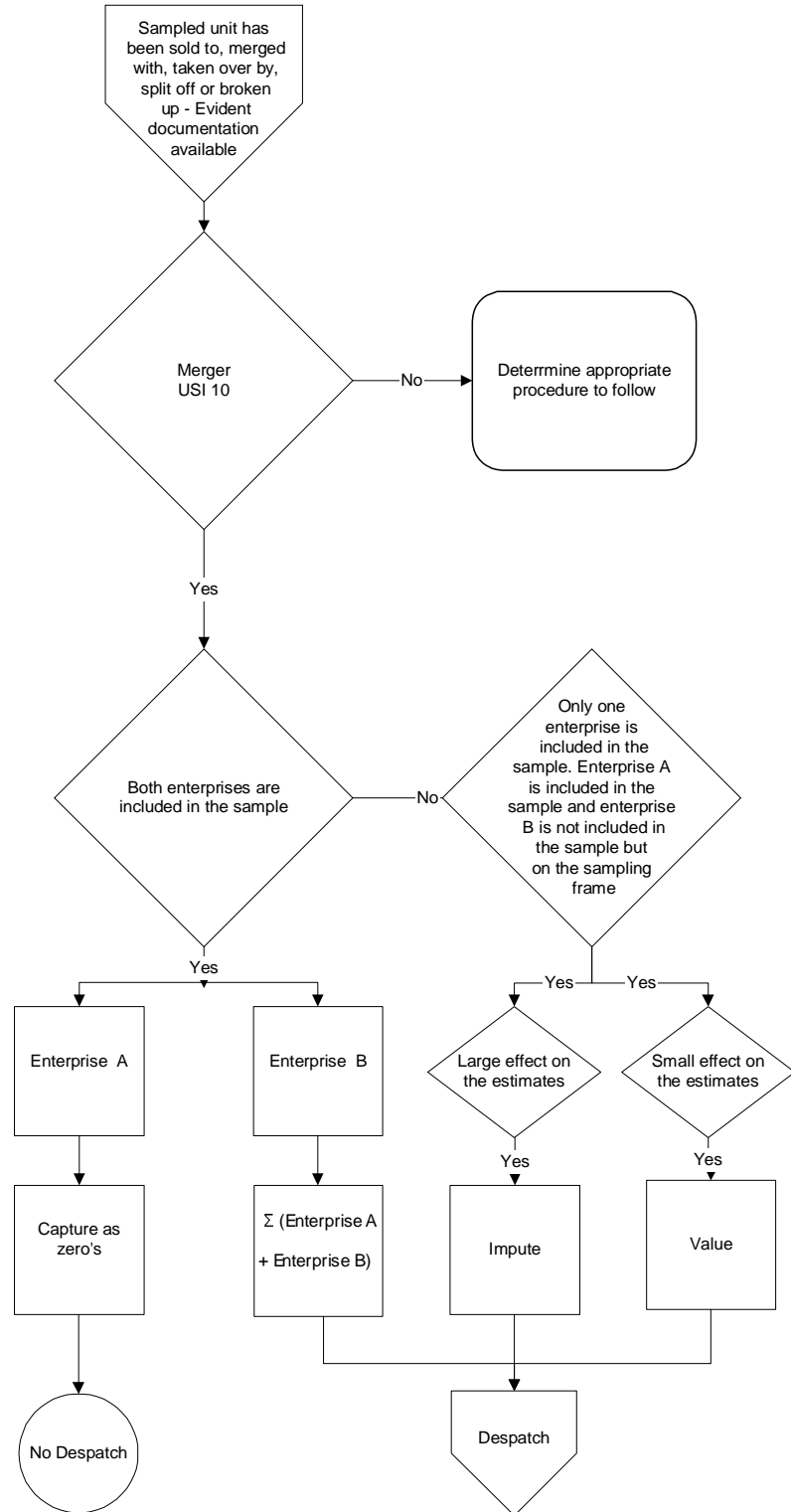
Note: If the unit cannot provide information on the enterprise level as requested and the only treatment is to impute the information then survey areas could decide whether they would re-dispatch questionnaires to these units.

2.3 Flowcharts of procedures in the sample maintenance in Economic Statistics

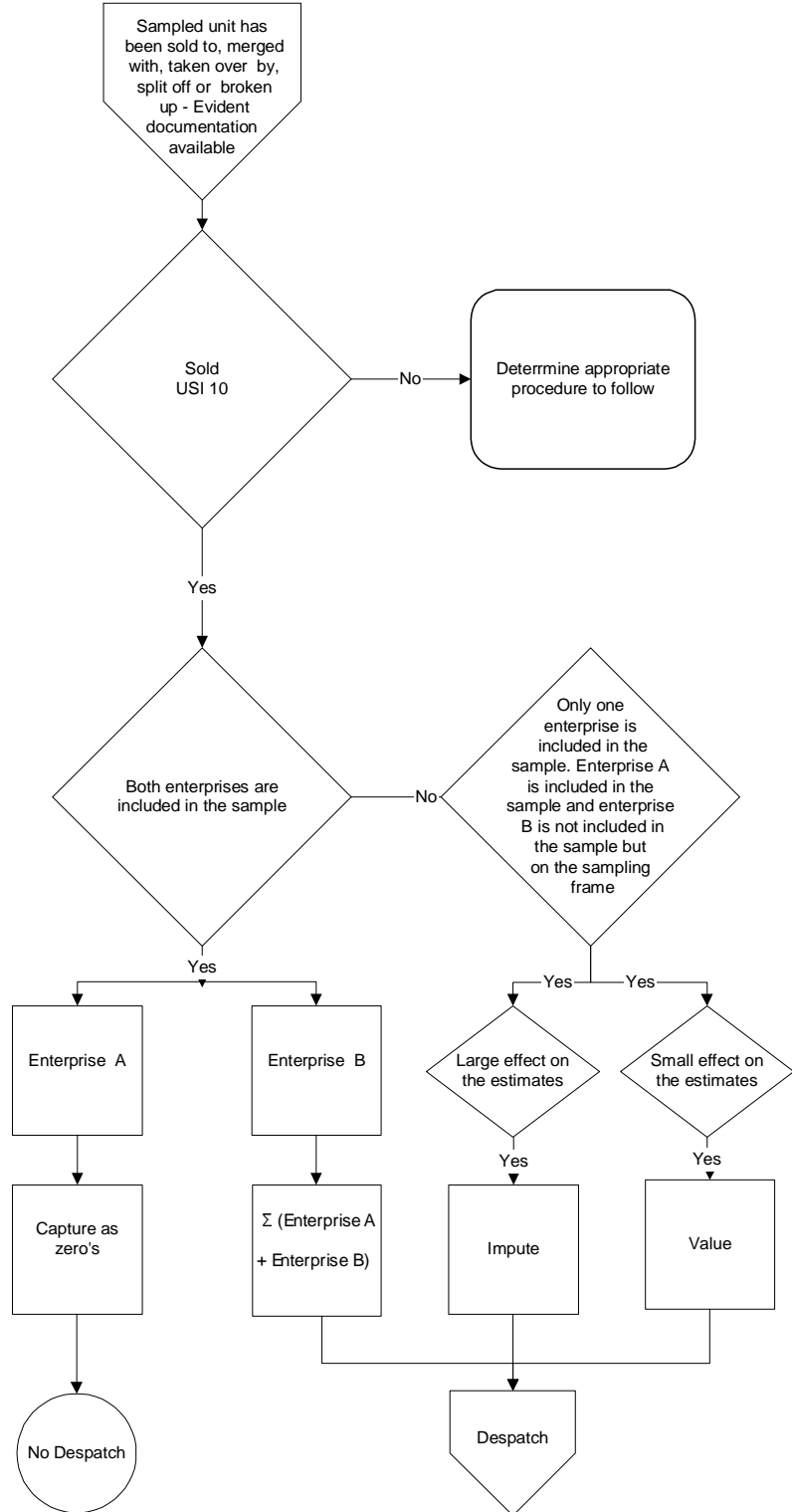
Break up



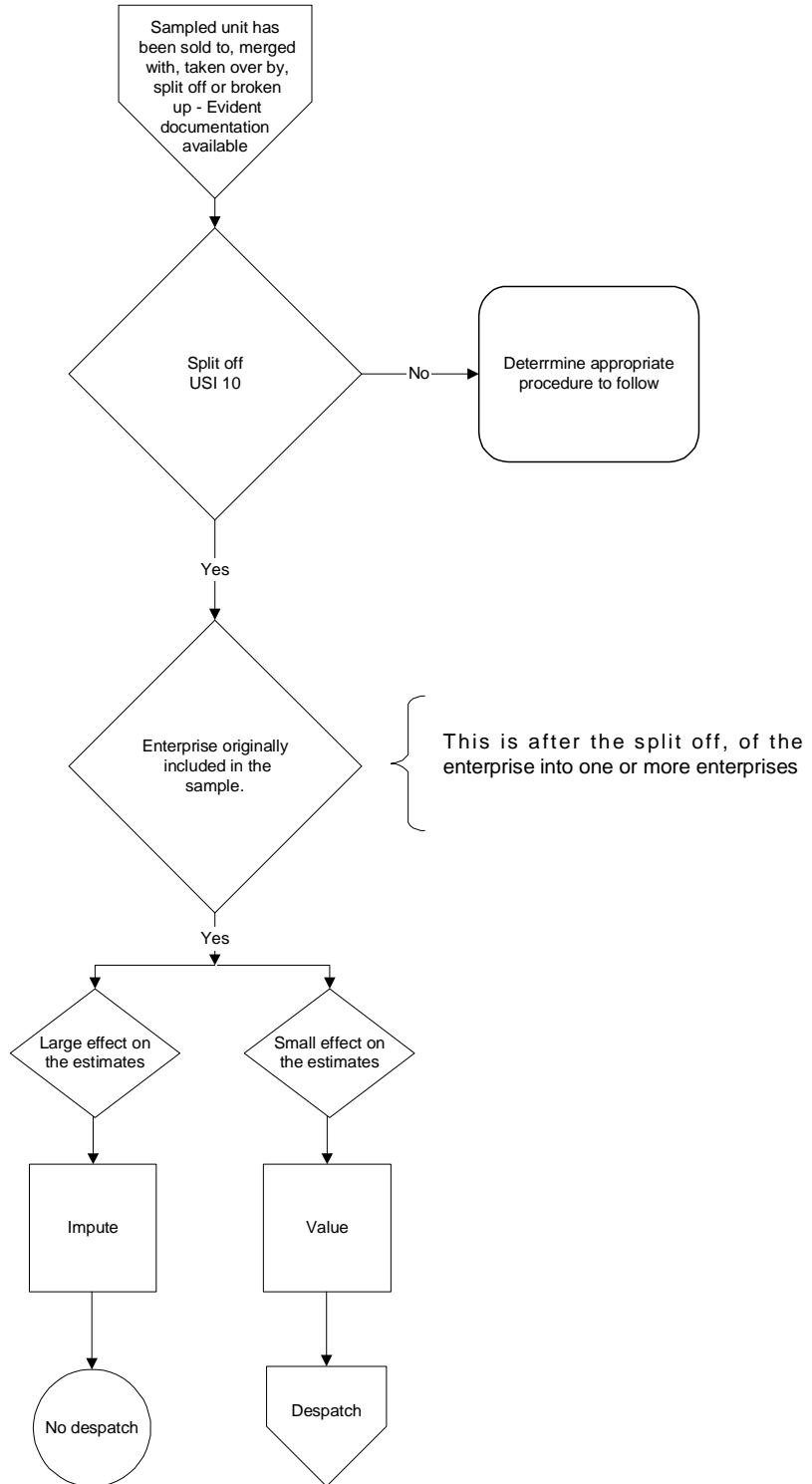
Merger



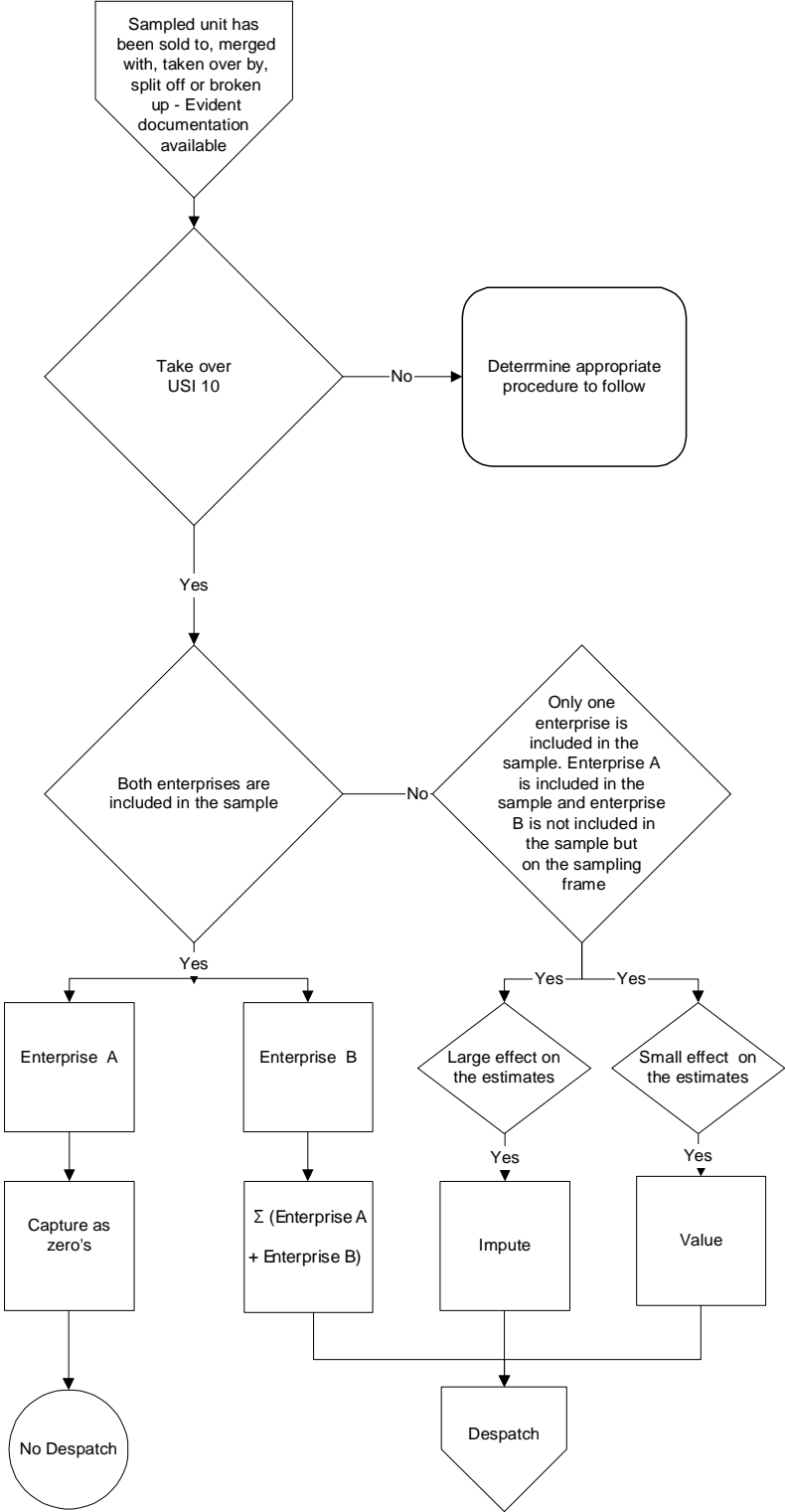
Sold



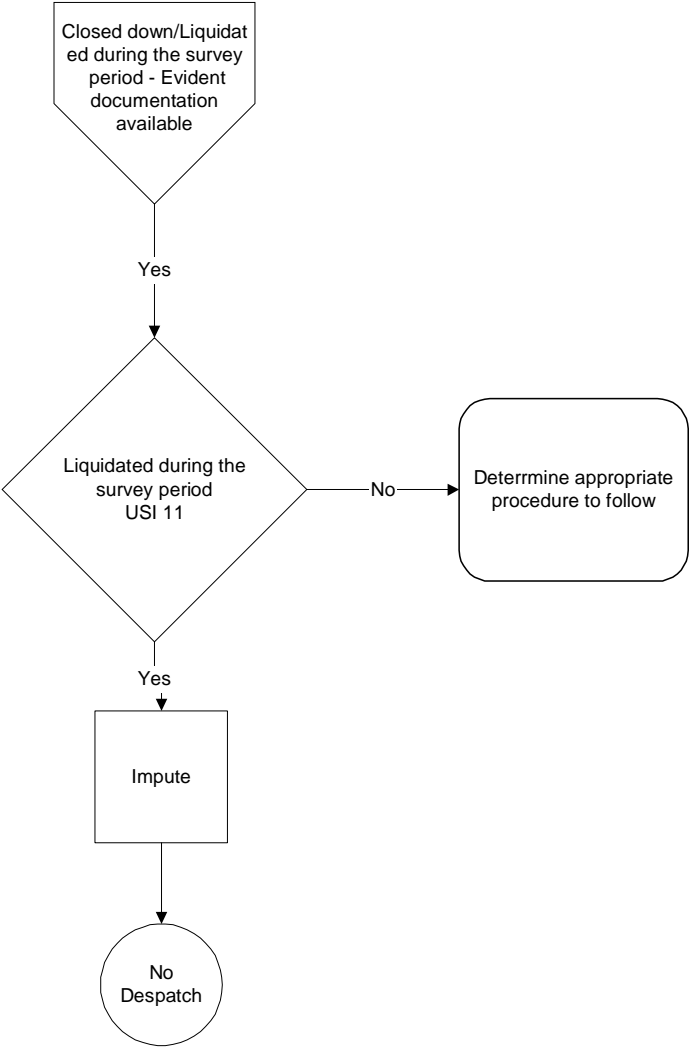
Split off



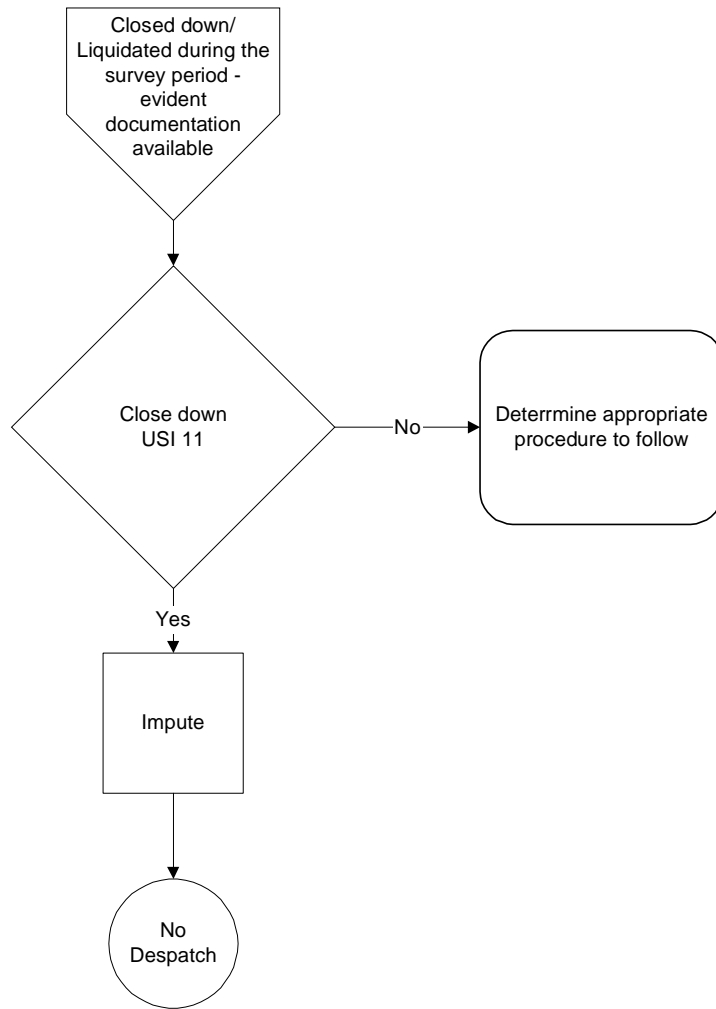
Take Over



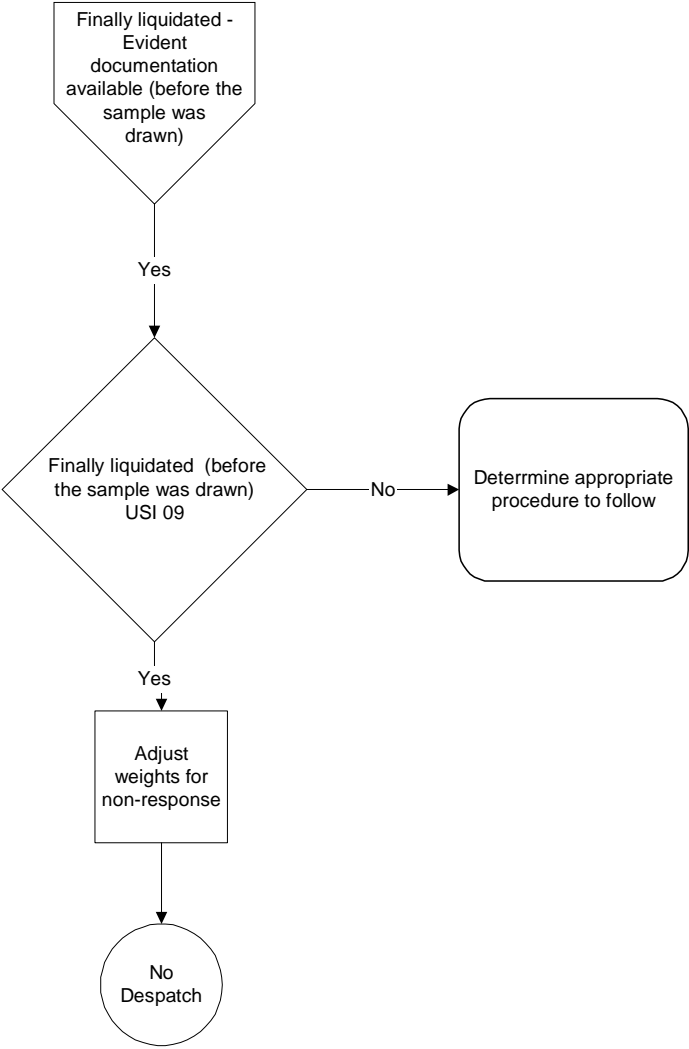
Liquidated during the survey period



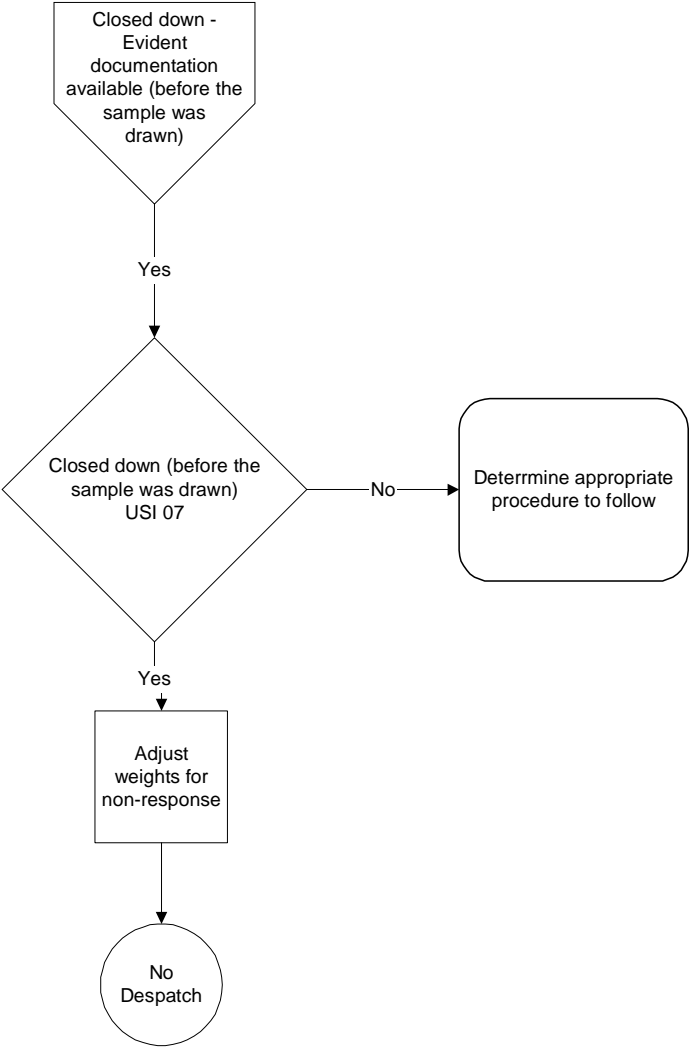
Closed down



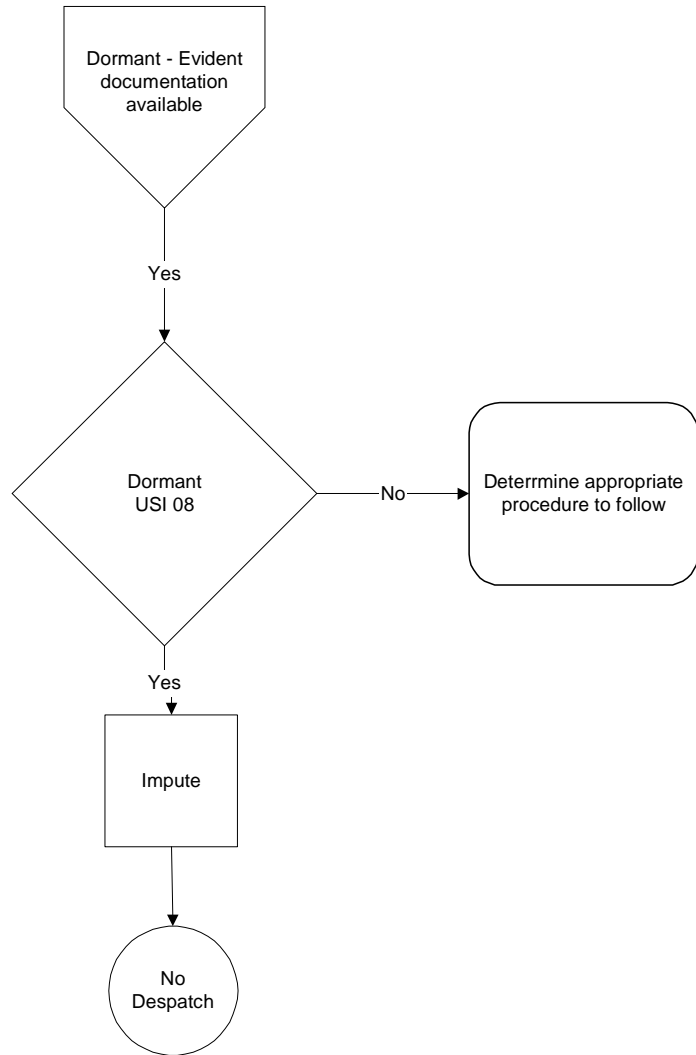
Finally liquidated (before the sample was drawn)



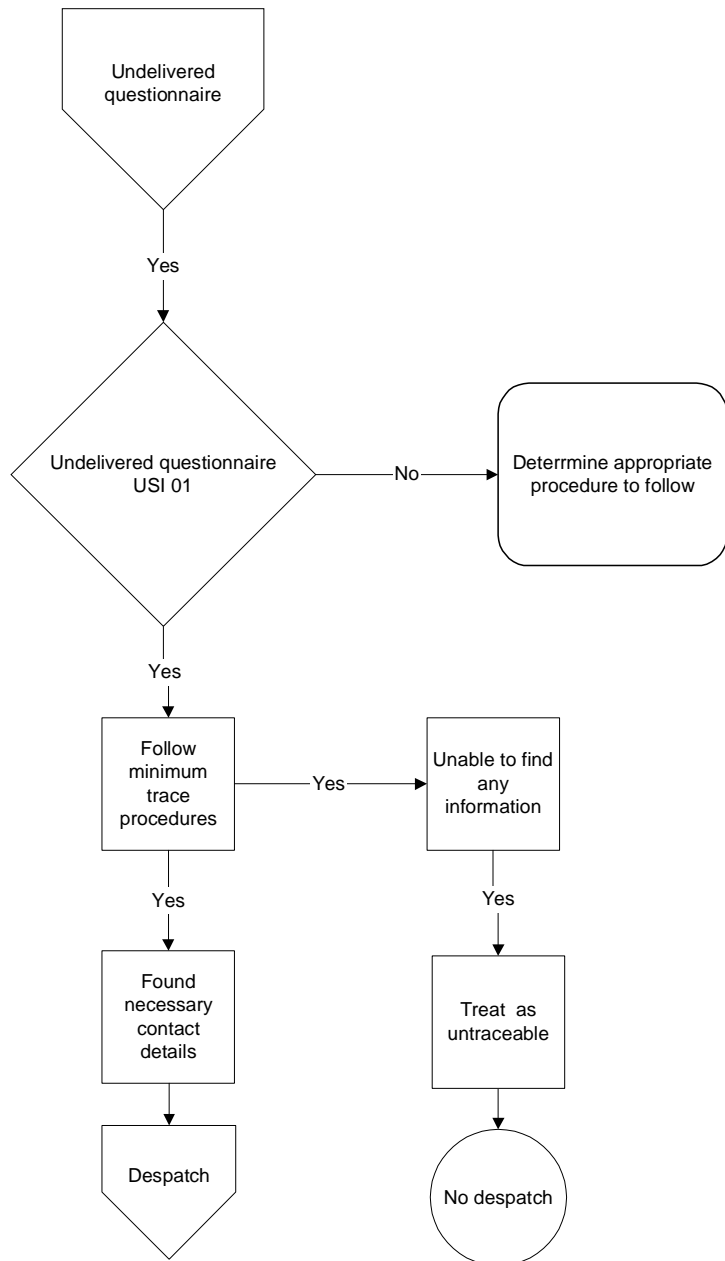
Closed down (before the sample was drawn)



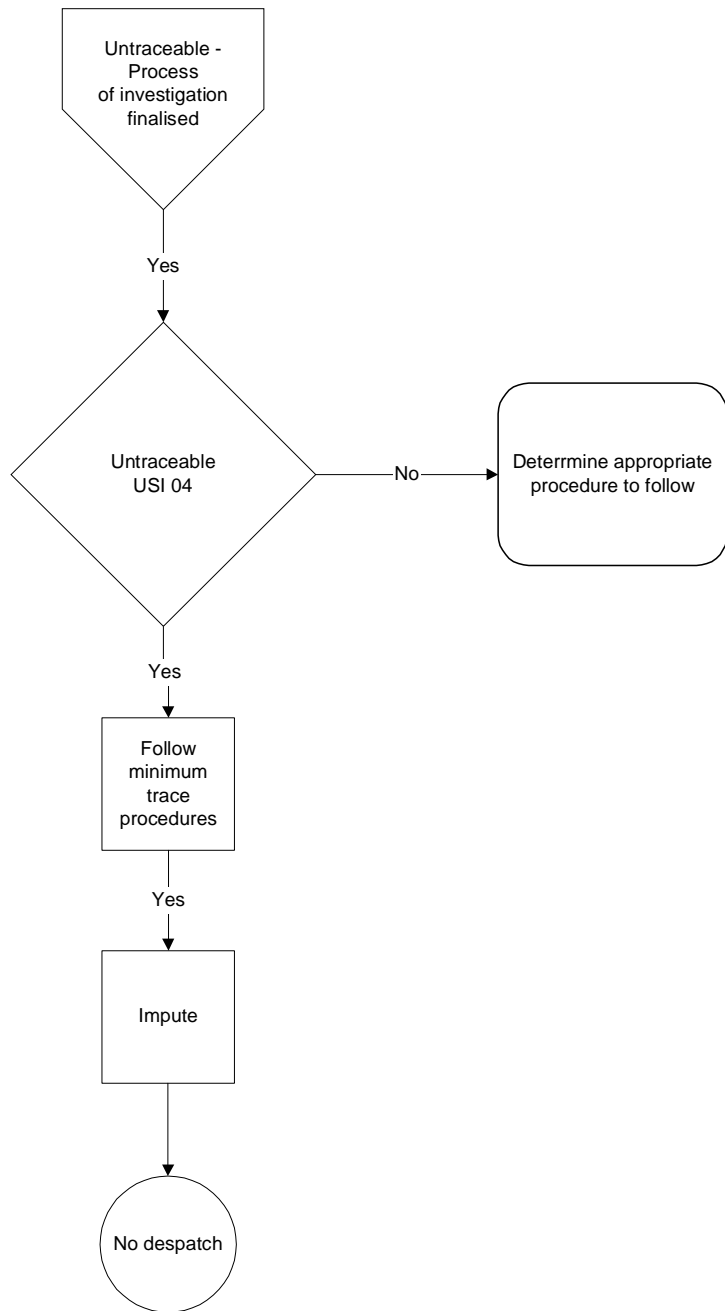
Dormant



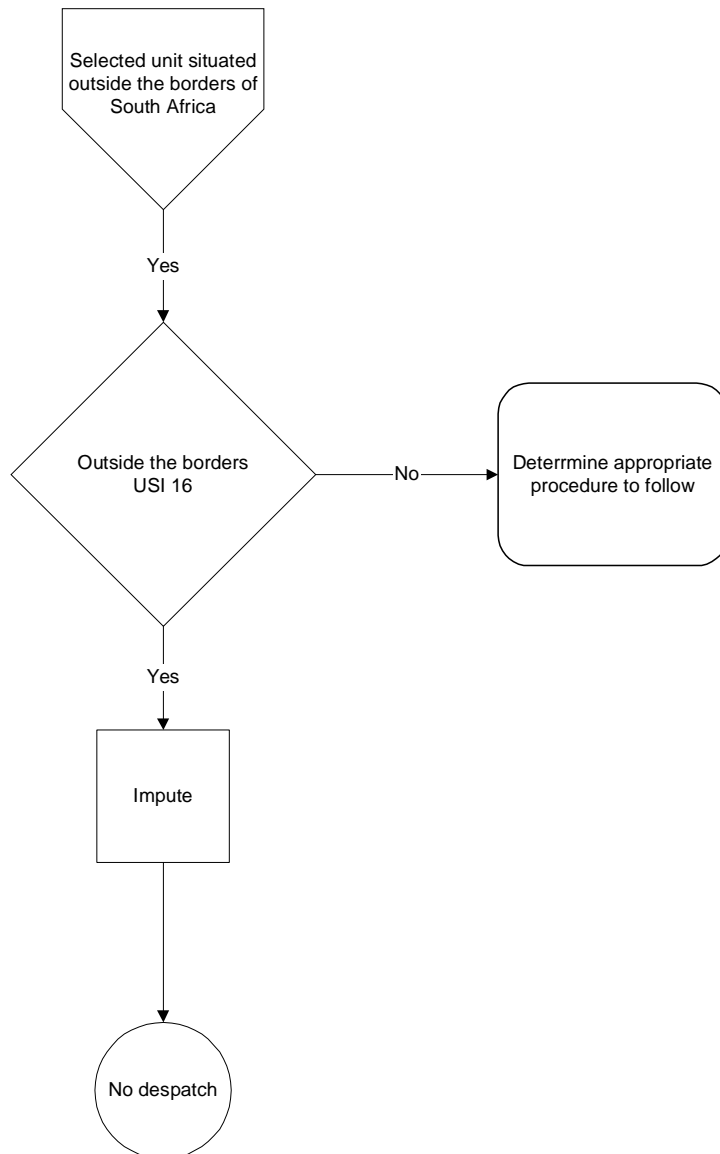
Undelivered questionnaire



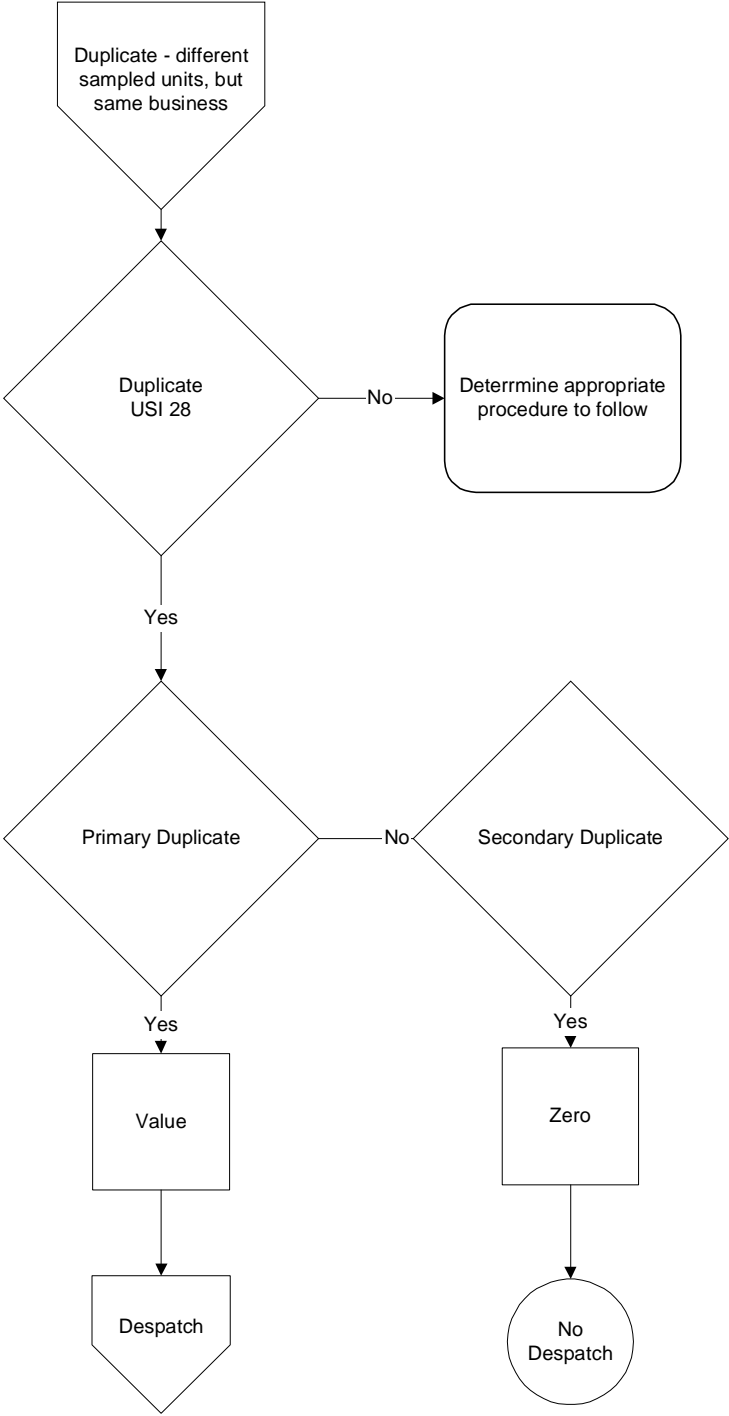
Untraceable



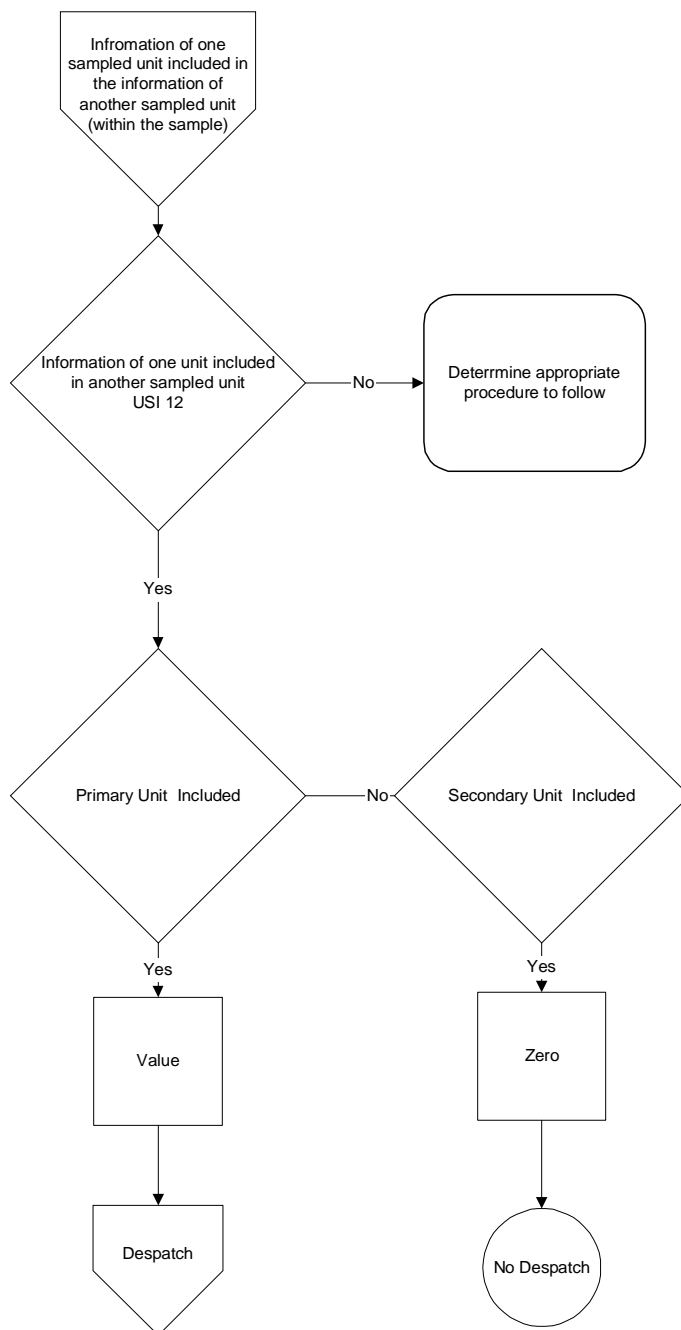
Outside the Borders of South Africa



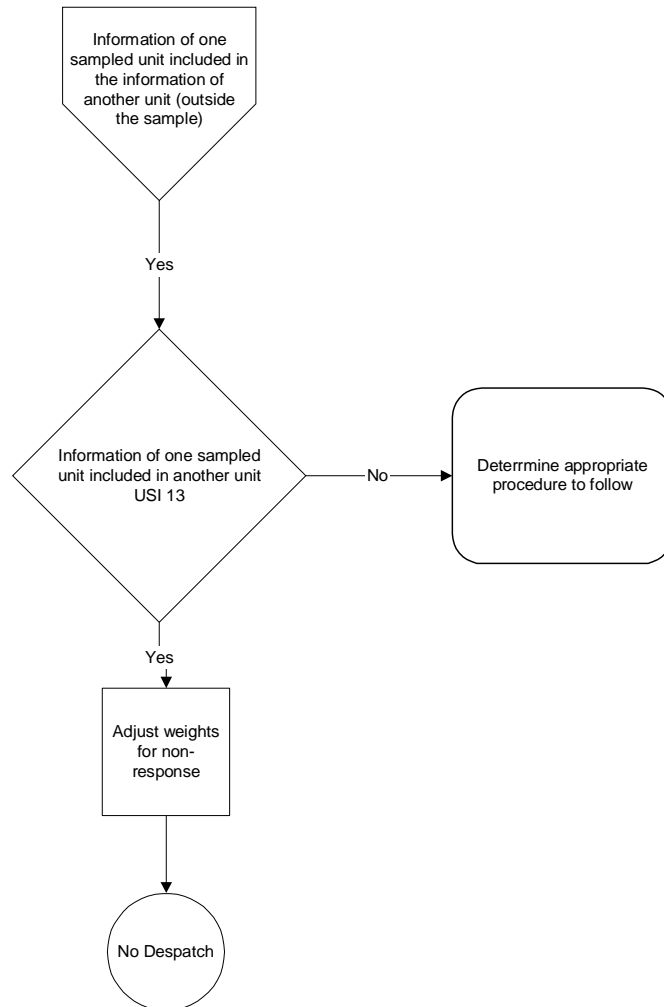
Duplicate



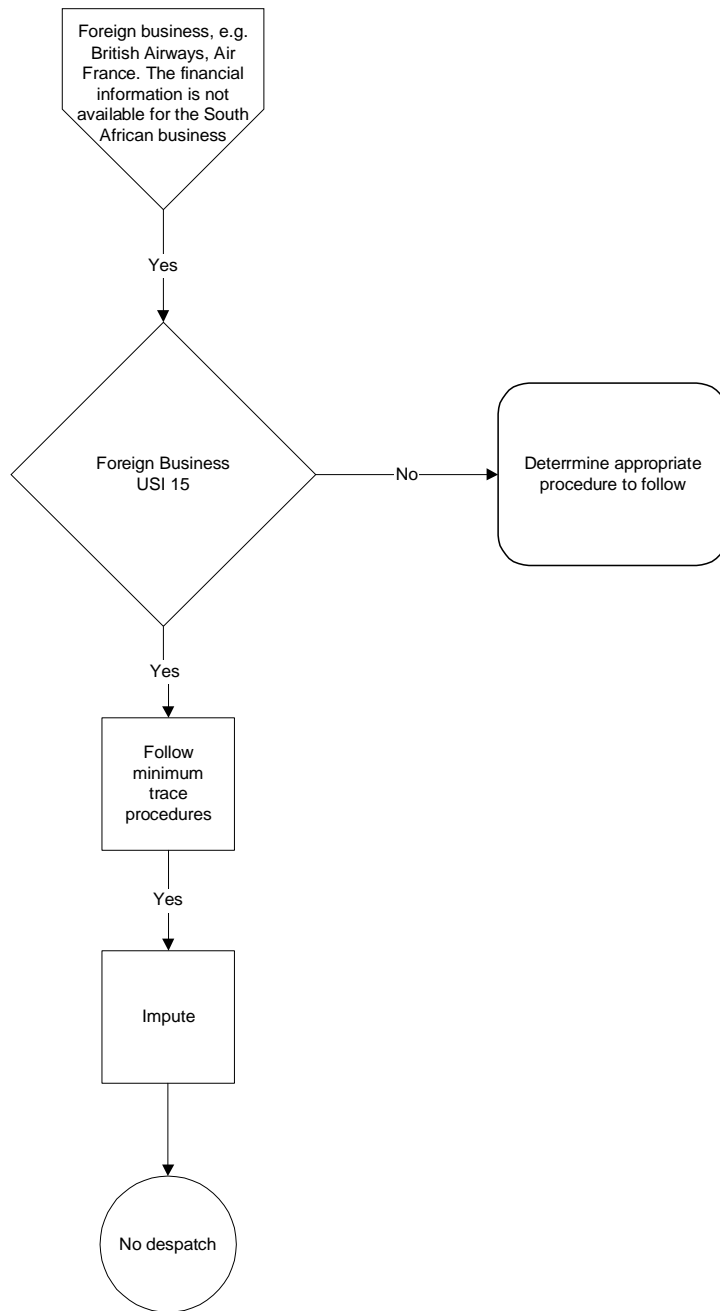
Information of one unit included in another sampled unit



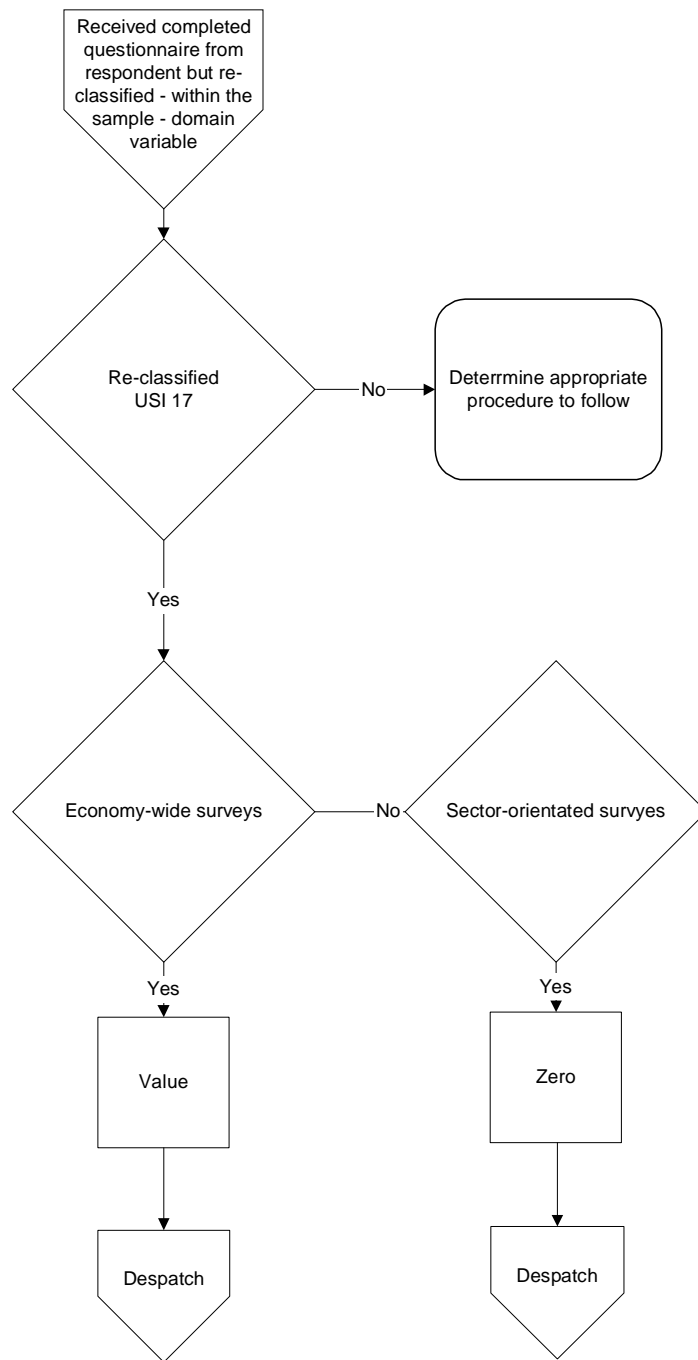
Information of one sampled unit included in another unit (outside the sample)



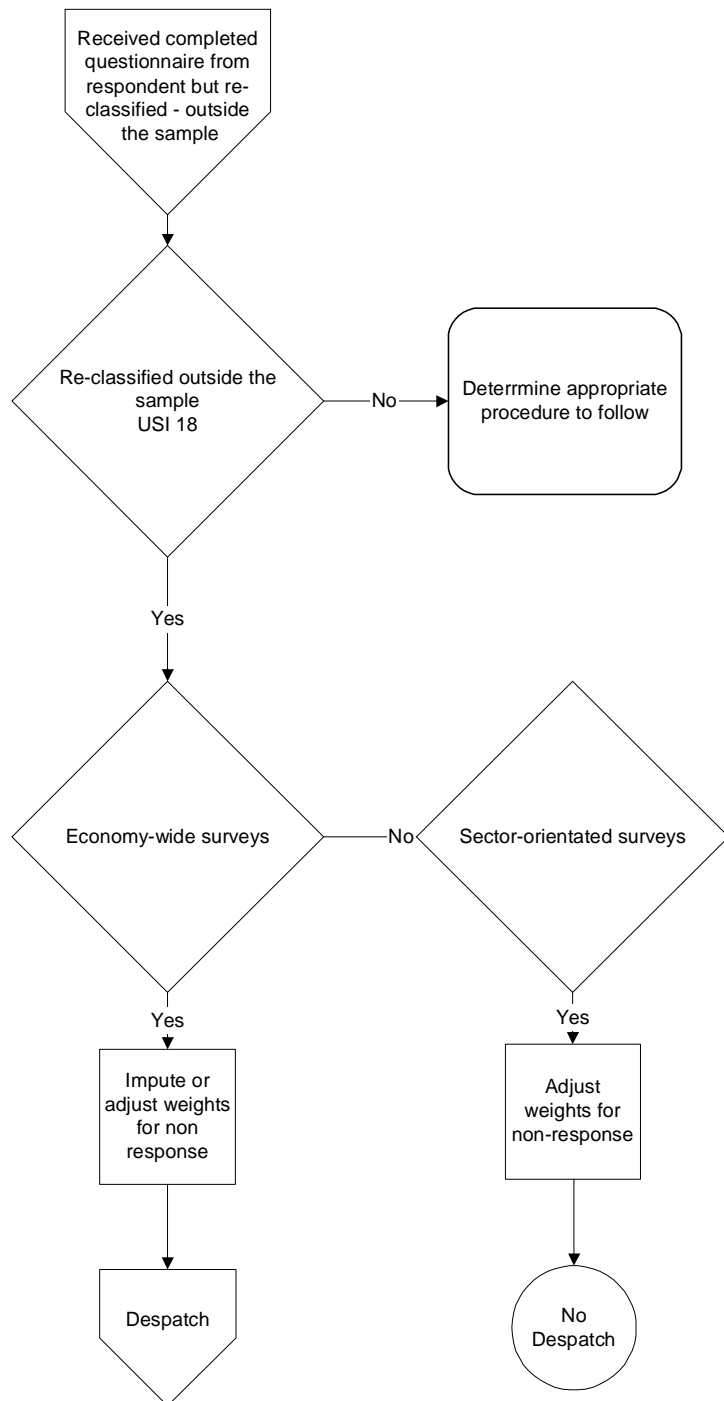
Foreign business



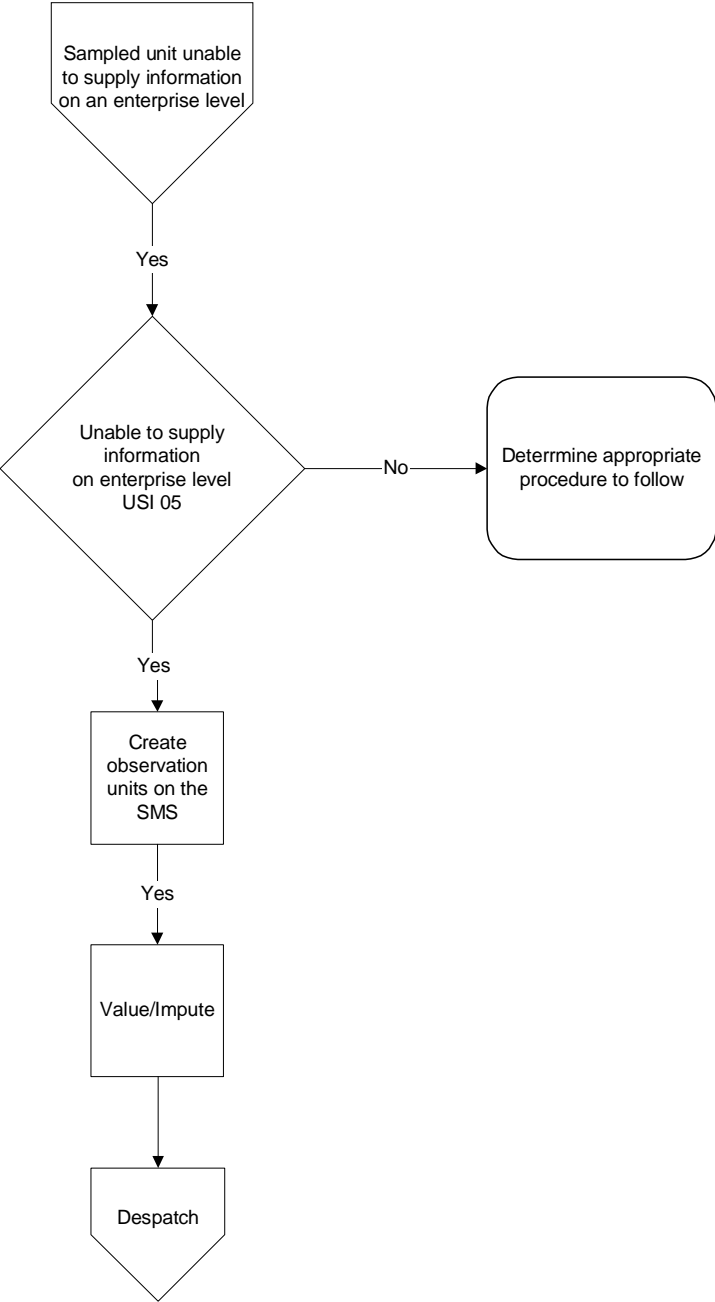
Re-classified - within the sample - domain variable



Re-classified - outside the sample



Unable to supply information on an enterprise level



General

If an enterprise has a return greater than R 300 000 when sampled and reported R 0 turnover then in this case, if we take the turnover as zero we cannot overestimate. We keep the enterprise with a zero turnover in the sample. For size group 2, 3 and 4 where the enterprises has a zero turnover, we can impute or if nothing can be done we treat it as non-response.

It is important to try to get the information from all the selected units in the sample; otherwise the appropriate outlined measure should be applied to the unit. All stratum 1 cases should be collected. In exceptional cases the Survey Statistician can impute the information by using the following methods: Regression imputation techniques, multiple imputation techniques or other techniques.

CLAN only takes into account the responding units. CLAN takes the number of respondents per stratum and the total number in the population and calculates a new weight for the subgroup or group depending on the level it was specified. CLAN will not run if there is no response within a stratum.

SIC classification could be combined if the response rate is very low (e.g. less than 50% or less than 10 respondents per size group). However, the classification groups should be very similar e.g. combine 611 and 612.

Whenever figures are released it would be expected that they are final figures and no changes would be made to the published figures. It could be assumed that the over-estimation and the under-estimation in a specific survey period would balance each other out and thus under this assumption we could expect that certain variables that were imputed for this period and that responded in this period would not be directly comparable. If a survey area is doing comparisons between two quarters etc. then the imputed units should not be compared with an actual response, because of the large fluctuation between the two records. If the unit was imputed in a previous quarter then the information that is received in the next quarter should be accepted without any further investigation of the unit. The respondent should not be requested to complete a questionnaire for the quarter that he/she didn't respond. This is totally unacceptable, this will also assist Stats SA to decrease the response burden of certain respondents.

CHAPTER 3

ESTIMATION

3.1 Integration of SMP into surveys

The standard Unit Status Indicator (USI) code has been introduced to consistently record the SMP treatment of units in estimation, future dispatch and feedback to the Business Sampling Frame as an indicator of unit changes. The function of the USI is to reflect the status of the unit at the time of collection and has no role in indicating how a unit has fared during editing.

The SMP treatment provide facilities to ensure that:

- current best practices are implemented consistently, cheaply and easily across the organisation;
- maintenance of the systems is much easier and cheaper; and
- changes in best practices can easily be reflected in the systems.

The USI is integrated as follows:

- The generalised software available to survey areas for recording USIs.
- Generalised estimation facilities should be developed for economic surveys. Surveys using these facilities must utilise the standard USI codes (response) as they are critical inputs into these estimation facilities. Other best practices for different processes in the survey cycle such as imputation and outliers, input editing/significance editing should also be developed for inclusion in the generalised facilities.

3.2 USI facilities

The facilities that have been developed for recording USIs include:

- selection of the appropriate response codes by processing staff.
- carrying forward of codes from the current cycle to a future survey cycle.

The selection of response codes is governed by a set of rules that indicate logically valid sets of codes.

3.3 Non-response

Definition

A non-respondent is a selected unit which is known to be 'live' but:

- is unable to report for all of the required data items;
- has refused to participate in the survey; or

- could not provide the data in time to be processed.

There are three main types of non-response:

- a. **Complete non-response** - the unit is known to be live, but the collection phase of the survey has been completed and no data has been obtained. These units may have been identified from situations where valid contact information was found but the respondent did not complete the survey questionnaire.
- b. **Refusal** - the unit is known to be live, but no data could be obtained as the respondent refused outright to participate in the survey. A refusal should always be referred to the responsible Survey Statistician.
- c. **Partial non-response** - the unit provides data for a subset of the data items on the questionnaire.

3.4 Imputation

Definition

Imputation involves the replacement of either missing or invalid data for a selected unit with acceptable data. It must be used in those cases where the selected unit has not responded or changes occurred. The extent to which data are imputed depends upon the amount and quality of data already provided. Imputation is used when data has not been provided, or the respondent has refused to participate in the survey. In some circumstances a survey area may decide to use complete imputation even if a partial response has been obtained.

Imputation methods

It is clear that imputed values will not always be the true values, but they need only be good enough for the purposes of providing outputs of acceptable quality. Imputation will usually involve the creation of values that are consistent with ideas about how the data should behave. Knowledge about the behaviour of the data should be obtained through a combination of subject matter expertise and analysis of relevant data. Imputation methods should be decided upon only after consultation with Quality and Methodology staff. Uniform guidelines for imputation are currently being developed by Quality and Methodology for implementation in a generalised imputation system.

There are two sources of data used in the imputation process.

- a. The first source is data provided by the selected unit itself. This data may come from a previous survey cycle as historical data or the current cycle if the unit was a partial non-respondent and has already provided some data. Surveys run on an infrequent or one off basis will not be able to use historical data in imputation.
- b. The second source is data from other units selected in the survey. This data may come from selected units in the current or a previous survey cycle. The most common method of imputation which utilises this source of data is the respondent mean imputation.

Respondent mean imputation involves generating a mean value for all missing data items from other units within the stratum for the entire reference period. These units may have reported data or had their values imputed reliably (e.g. imputes from previous data provide a good estimate of that unit's current value). The respondent mean includes units which are still operating even though they may contribute zero to each data item on the questionnaire (e.g. secondary duplicate, out of scope, etc.), as it is possible that the non-responding unit is also one of these units.

Another method of imputation, which is more likely to be used for larger businesses, might be to take reported data from a donor unit that has similar characteristics to the unit undergoing imputation.

Any type of imputation performed using only data from other units selected in the current cycle should not contribute to estimates of variance (e.g. mean or donor impute). These imputes must be distinguished from other types of imputes which should contribute to estimates of variance.

Relevant USI response codes

Descriptions for Version 2 of the Survey Management System (SMS).

USI code	Description
00	Outstanding questionnaire
01	Undelivered questionnaire
02	Undelivered reminder
03	Enterprise under investigation
04	Untraceable – Process of investigation finalised
05	Sampled unit unable to supply information on an enterprise level ⁵
06	Extension after due date
07	Closed down – Evident documentation available (before the sample was drawn)
08	Dormant – Evident documentation available
09	Finally liquidated – Evident documentation available ⁶ (before the sample was drawn)

⁵ Create observation unit on the SMS.

USI code	Description
10	Sampled unit has been sold to, merged with, taken over by, split-off or broken up – Evident documentation available
11	Closed down/Liquidated during the survey period – Evident documentation available
12	Information of one sampled unit included in the information of another sampled unit (within the sample)
13	Information of one sampled unit included in the information of another unit (outside the sample)
14	Received completed questionnaire from respondent with no employees working in the sampled unit, e.g. Property owning and letting companies (SIC 84110)
15	Foreign businesses, e.g. British Airways, Air France. The financial information is not available for the South Africa business ⁷
16	Selected unit situated outside the borders of South Africa ⁸
17	Received completed questionnaire from respondent but re-classified – Within the sample - domain variable
18	Received completed questionnaire from respondent but re-classified – Outside the sample ⁴
19	Only a Part 1 or a Part 2 questionnaire received, respondent to be followed-up (SEOGR survey)
20	Legal action to be taken
21	Manual imputation
22	Received completed questionnaire from respondent
23	Received financial statement or pay-roll printouts from respondent
24	Omit the sampled unit for a short period, e.g. computer systems currently not working, a fire broke out in the factory, etc.
25	Checked questionnaire

⁶ If still in the process of liquidation the completed questionnaire must be supplied to Stats SA. The USI code will reflect '00' until receipt of the completed questionnaire, whereafter the USI code will reflect '22'.

⁷ A questionnaire must not be sent in the next survey period. Only applicable to Financial Statistics.

⁸ Out of scope.

USI code	Description
26	Edited questionnaire
27	Edited financial statement
28	Duplicate – different sampled units, but same business
29	ASO refers investigation to supervisor

Note – Changes to name, address and other contact details are to be changed on the SMS. Before any name and/or address changes are done the collection clerk should phone the respondent and enquire if this name and/or address change did not occur because of a demographic event. A demographic event includes that of sold, merged, taken-over, split-off and break-up i.e. USI code 10. If not a demographic event then the collection clerk must allocate the applicable USI code according to the new information received from the respondent.

If the only change is a name change then during the next round of surveying the questionnaire must be posted with the new name and/or address. It is important to remember that a USI code cannot be allocated for a change in name and/or address.

3.5 Estimation

Surveys are used as a tool to collect information from some or all units of a population in order to answer a question or make a decision. We would like to make inferences on the values of some population characteristics based on information obtained from a sample. To infer about a population characteristic we normally estimate its value using observations from a sample. This process is called estimation. In a sample survey estimates for the entire population are derived by expanding or weighting the sample data.

Only those units which have reached an "acceptable" level of "cleanliness" during data editing may be used to produce estimates.

3.6 Outliers

An outlier is an observation that has a major effect on an estimate. Because of its independently known atypical nature, it needs special treatment. The correctness of the values in the observation is not being questioned. Outliers are usually discovered during the output-editing phase of the survey cycle. Before choosing the treatment for an outlier it is essential to know the reason for its occurrence and whether the outlier provides any information about the population of interest. These units cannot be excluded from the estimates since they are part of the population and provide some information about the

population. If nothing else, the units provide information about themselves. The crucial question is whether the unit is representative of any other units in the same stratum.

Outliers can belong to one or more categories. Firstly, there are outliers in level which occur when level estimates are of interest. These outliers are normally unusually large values. Secondly there are outliers in time, when a unit takes an unusual value relative to its past behaviour.

3.7 Other systems issues

Collection of auxiliary information

Under each of the data collection procedures the storing of extra information that has been collected on the survey system has been mentioned. This information may be needed for estimation, future dispatch, monitoring purposes.

For each unit the following fields need to be collected for future dispatch purposes:

- legal entity name;
- geographical information/location name (if applicable);
- trading name; and
- mailing address.

The following fields, if collected by the survey, need to be stored for domain analysis in the estimation process:

- reported SIC; and
- reported employment.

CHAPTER 4

GENERAL COMMENTS

4.1 Introduction

This chapter provides the justifications of the individual data collection procedures described in Chapter 2, Data collection procedures, of this manual. The intended audiences for this chapter are staff reviewing the procedures, and staff (from survey areas) dealing with the more difficult situations who need to understand the justifications behind the procedures.

Data collection procedures are necessary to assist staff in identifying and treating differences between units on the sample and in the real world. There are wide ranges of possible differences that can be encountered in Stats SA's economic surveys. If not correctly treated by the survey areas some of these differences will result in large errors in survey estimates. To minimise these errors, it is important to ensure that the differences are identified and treated correctly and consistently across the entire range of Stats SA's economic surveys.

The procedures given in this manual must be simple and practical to implement consistently across economic surveys. The simplicity of the procedures will allow staff of varying levels of experience to deal with them and achieve consistent solutions. The treatment of each type of difference should be such that it is simple to implement consistently, duplication is avoided, coverage is maximised, representation in the population is appropriate, and action for future dispatch is appropriate.

Procedures in this manual do not attempt to prescribe treatments for every possible difference that may occur. Sample maintenance procedures give simple rules and criteria for the most common types of differences. Many complex queries will not fall within the scope of the procedures given in this manual. These cases should be dealt with individually by Quality and Methodology in collaboration with the specific survey area, using the concepts outlined below.

In most instances, the procedures were designed to be robust against a change in environment. The procedures should not change regardless of the maintenance procedures for the Business Sampling Frame. In the event of a large change to an environment, however, some criteria will need to be adjusted.

Section 4.2 contains the justification of each current data collection procedure and includes some discussion as to why the particular methods of identification and treatment of differences were chosen. Where relevant this discussion includes reasons for rejecting other options.

4.2 Justification for the current data collection procedures

Normal unit procedures apply to selected units, which do not require any special treatment in data collection, estimation or future dispatch (all these changes would be made on the SMS). It is possible that there exist differences between the description of the selected unit and real world business. An example of the differences is changes/corrections to basic contact information. Although it is possible to code each of these different types of normal units separately, only one category was used due to:

1. The simplicity and reliability improvement in having fewer categories available to choose from; and
2. Information produced from a finer breakdown of common differences not being meaningful in all situations. For example, suppose that fine coding was available, and the only difference for business A was a change of contact address, whilst business B was involved in a takeover of operations and also had a contact address change. Since businesses can only be assigned to one category or USI code for treatment purposes, business A would be assigned to the "change of contact address" category and business B would most likely be assigned to the structural change category (as a takeover takes priority over an address change due to the effect on estimation, etc.). Coding to such a fine level of detail like "change of contact address" becomes meaningless.

Note: Name changes will be recorded only on the SMS and no USI code exist for the use of name changes.

The deathing rule is based on the presence of income and/or expenses for a business. The rationale for this is that these attributes are unambiguous and can be consistently implemented. A business cannot be considered as live and operating without income and/or expense attributes; and respondents will have information on whether or not the business generated any income or incurred any expenses during the reference period.

There are many other alternatives in identifying businesses that have ceased operating that could also be used by survey areas. Some of these alternatives include one or more of the following criteria: the respondent told Stats SA that the business has ceased operating; the business has no assets or liabilities; all locations for the business have been closed down; the legal entity for the business has been de-registered with the SARS; etc.

A written letter for the reasons given below should confirm these options from the deathing criteria.

1. Experience has shown that respondents may think their business has ceased because they've closed down part of the operations collected in a particular survey, or they've ceased employing, or closed down locations, etc. Consistency

- across surveys should be achieved, as there are many different interpretations that can be made by both the respondent and the Stats SA officer.
2. Activities such as de-registering with the SARS do not always occur as soon as a business has ceased generating income or incurring expenses. This lag is mainly due to delays in notifying the appropriate office and in the processing of that notification. Any business, which is still generating income or incurring expenses, is still considered to be in scope of the sample.
 3. The closing down of locations could not, on that basis only, form part of the deathing criteria since it is possible to operate a business without a specific location (although in most cases it would imply that a business has ceased operating).

The intention of the business to operate in the future was included in the deathing criteria to differentiate those businesses that have permanently ceased operation (confirmed deaths where there would be no further dispatch of questionnaires) from those which have only temporarily ceased operation (where further dispatching of questionnaires would be required).

To avoid inappropriate deathing, confirmation that a business has ceased should be obtained from a reliable source. Reliable sources include contact with a director, employee, partner, accountant, liquidator, or receiver of the business. Confirmation from someone who bought the business or a next door neighbour would also be considered as a source of information and contact information should still be sought.

Structural changes to a selected unit appear when the business corresponding to the selected unit's identification name has ceased operating **and** some or all of the operations of the business were transferred, sold, taken over and/or merged with the operations of a real world business which is either in scope of the survey or not.

There are two types of duplications that most commonly occur. There is duplication caused by real world changes (such as sales, takeovers, mergers) – these are covered under the structural change procedures – and duplication where no real world change occurred (such as spelling errors in the frame identification names) and these are covered under the duplicate procedures.

The sampled units are sufficient to represent the whole population, including the primary duplicate. Note that the SMP only correct for bias caused by duplicates in the sample by treating the duplicates that are actually identified, the number of which is expected to be small for samples sourced from the Business Sampling Frame. It would be unusual to select (and therefore identify) duplicated units unless the survey had a high sampling fraction. Therefore the sample may not accurately reflect the proportion of duplicates on the frame. If a large difference between the proportion of duplicates in the sample and on the frame is suspected, more complex procedures may be required for treatment.

Primary duplicates are considered to be live units for the purposes of the survey. All outstanding units at the end of survey processing are assumed to be 'live'. It is therefore

possible for non-responding units requiring imputation to also be duplicates. That is because since the businesses haven't responded Stats SA does not know whether or not they received two questionnaires for the same business.

NOTE: Only one USI code can be assigned to a selected unit.

Selected units found to be temporarily non-employing or reporting zero employment are in scope of sample-based surveys unless non-employers are excluded from the survey scope. While a business continues to generate income or incur expenses it continues to remain in scope of the survey, even if it ceases to employ staff.

4.3 Minimum procedures

The minimum set of steps to follow for undelivered post or untraceable units and Intensive Follow Up of units are listed below. The order of these steps is intended as a guide and it may vary depending on the individual circumstances of the unit and any additional information that may be available for it. It may also be necessary to repeat steps if additional information is found from other sources.

The purpose of these steps is to find alternative contact information for businesses. If a phone number was found then try to contact the respondent at least three different times and over different days spread over 1-2 weeks. If alternative address (es) was (were) found, dispatch to the alternative address (es).














Minimum procedures:



1. Check the Business Sampling Frame and the SMS for other survey area's contact information etc.
2. Phone 1023
3. Phone 10118 or use yellow pages on the internet
4. Use the telephone directory (hard copy)
5. Make use of the DTI database
6. Make use of McGregor's database.
7. Fax query through to Telkom Customer Service, where a search will be done on the requested information.

Note that these are the minimum unable to trace procedures and where appropriate, other sources may be used in addition of the above.

4.4 Response rates

What should be counted as an actual response in the survey areas?

USI code	Description	Should be counted as a response	Methodological treatment
00	Outstanding questionnaire	No	Imputed Value
01	Undelivered questionnaire	No	See 
02	Undelivered reminder	No	Imputed Value
03	Enterprise under investigation	No	Imputed Value
04	Untraceable – Process of investigation finalised	No	See 
05	Sampled unit unable to supply information on an enterprise level	No	See 
06	Extension after due date	No	Imputed Value
07	Closed down – Evident documentation available (before the sample was drawn)	No	See 
08	Dormant – Evident documentation available	No	See 
09	Finally liquidated – Evident documentation available (before the sample was drawn)	No	See 
10	Sampled unit has been sold to, merged with, taken over by, split-off or broken up – Evident documentation available	No	See 
11	Closed down/Liquidated during the survey period – Evident documentation available	No	See 
12	Information of one sampled unit included in the information of another sampled unit (within the sample)	No	See 
13	Information of one sampled unit included in the information of another unit (outside the sample)	No	See 
14	Received completed questionnaire from respondent with no employees working in the sampled unit, e.g. Property owning and letting companies (SIC 84110)	Yes	Actual Value
15	Foreign businesses, e.g. British Airways, Air France. The financial information is not available for the South African business	No	See 
16	Selected unit situated outside the borders of South Africa	No	See 
17	Received completed questionnaire from	Yes	See 

	respondent but re-classified – Within the sample - domain variable		
18	Received completed questionnaire from respondent but re-classified – Outside the sample	No	See 
19	Only a Part 1 or a Part 2 questionnaire received, respondent to be followed up (SEOGR survey)	No	N/A
20	Legal action to be taken	No	Imputed Value
21	Manual imputation	No	Imputed Value
22	Received completed questionnaire from respondent	Yes	Actual Value
23	Received financial statement or pay-roll printouts from respondent	Yes	Actual Value
24	Omit the sampled unit for a short period, e.g. computer systems currently not working, a fire broke out in the factory, etc.	No	Imputed Value
25	Checked questionnaire	Yes	Actual Value
26	Edited questionnaire	Yes	Actual Value
27	Edited financial statement	Yes	Actual Value
28	Duplicate – different sampled units, but same business	Yes	See 
29	ASO refers investigation to supervisor	No	Imputed Value

CHAPTER 5

GLOSSARY

1. Definitions

Aggregation

The grouping of units into categories and the summing of values within these categories to produce macro data.

Business Sampling Frame

The Business Sampling Frame is essentially a list of all employing businesses operating in South Africa with an annual turnover greater than or equal to R 300 000. The prime purpose of the Business Sampling Frame is to provide a comprehensive source of business names and addresses from which selections can be made for inclusion in Stats SA's economic surveys. The Business Sampling Frame provides an integrated medium for recording the hierarchical structure of Statistical Units.

Business Sampling Frame ('Snapshot')

A business sampling frame ('Snapshot') refers to the regular yearly creation of a sample frame at the different levels of the statistical units model (enterprise, kind of activity and geographic unit) and the centralised checking of those frames to ensure their quality. The checked frame would then be released to the various collection areas ready for them to stratify the frame, select samples, questionnaire reporting units (if necessary), perform a set of best practice checks on the sample, and generally continue with the usual dispatch process.

Classification

A classification is an arrangement of populations of people or organisations or other types of units into classes according to the characteristics they have in common.

Collection unit

A collection unit is a unit from which data are collected.

Data

A representation of facts, concepts, or instructions in a formalised manner, suitable for communication, interpretation, or processing by humans or by automatic means.

Data capture

The process of transferring survey data from questionnaires to a machine-readable medium.

Enterprise

An enterprise is a legal unit or a combination of legal units that includes and, directly controls all functions necessary to carry out its production activities.

Estimate

Statistic calculated from the sample from which inferences are made about the entire population of interest. It may be a total, a proportion or a mean.

Imputation

Imputation involves the replacement of missing or invalid data for a selected unit with acceptable data. This only applies in the case of non-response.

Intensive Follow Up

Intensive Follow Up (IFU) occurs when a respondent has not provided any or insufficient information in response to a collection; it is the process of collecting the information from the respondent using the timeliest approach available. The approach that tends to be used for IFU is telephone interviewing, although facsimile is also used at times.

Liquidation

A company in liquidation is being wound up (i.e. cease to exist), its assets will be sold off and the proceeds distributed to creditors and/or the owners.

Non-response

A selected unit that is known to be live but is unable to report for all required data items, has refused to participate in the survey or could not provide the data in time to be processed.

Observation Unit

A unit for which data are reported that does not equate to the statistical unit on the sampling frame, i.e. an aggregation or dissection of statistical units.

Population

In the context of a collection, the term 'population' can be used to mean either the 'target population' or 'survey population' of a collection, although once a collection is underway it usually refers to the survey population. In the context of Stats SA's datasets, the population refers to the actual set of unit objects of interest which are presented in the dataset about which information can be inferred in relation to the dataset.

Reference period

The period of time for which the data are collected and/or compiled (e.g. calendar year, financial year, quarter month, specific day of month, etc.).

Respondent

The person/people which provide the information about the observation unit.

Sample

A part of a population selected for the purpose of studying certain characteristics of an entire population of interest.

Sampling frame

A list of all sampling units in the survey population available for selection at the time of sample selection. It should be comprehensive, complete and up-to-date to keep bias to a minimum.

Sampling unit

A statistical unit on the sampling frame which is available for selection.

Scope/Survey Population

The set of units about which information is required.

Statistical unit

A unit about which statistics are compiled.

Survey

The collection of information about characteristics of interest from some, or all, units of a population using well-defined concepts, methods and procedures, and the compilation of such information into a useful summary questionnaire.

Survey cycle

The specific instance in which a survey is conducted

Variance

A measure of the precision of the estimate that takes into account the fact that only a subset of the population has been sampled (rather than completely enumerated).

2. Acronyms

BCM	Business Classification and Maintenance
EN	Enterprise
IFU	Intensive Follow Up
SARS	South African Revenue Services
SIC	Standard Industrial Classification
SMP	Sample Maintenance Procedure
Stats SA	Statistics South Africa
USI	Unit Status Indicator
VAT	Value Added Tax

3. References

ABS sample and frame maintenance procedures reference manual, June 1999.

Weighting and re-weighting in CVTS2, Claes Andersson, Statistics Sweden, 2002.

Frame problems: Suggestions on treatments of some typical cases, Lennart Nordberg, Statistics Sweden, 2002.