



DIAGNOSTIC FRAMEWORK: National Accounts and Supporting Statistics

SELF ASSESSMENT TOOL

Data sources (surveys, compilation)

Seminar on developing the capacity to produce economic statistics,
including national accounts in accordance with the 2008 SNA, in the Asian
and Pacific region

10-13 October 2011, Bangkok, Thailand

Benson Sim

UN Statistics Division/ DESA

1



Outline of presentation

- **Diagnostic Framework for National Accounts and Supporting Economic Statistics (DF-NA&ES)**
- Elaborating on elements of the DF tool relating to (for this session)
 - **Data Sources – surveys (4.3.3)**
 - Household income and expenditure survey
 - Labour force surveys (case of 1-2 survey)
 - Enterprise surveys – case of industrial statistics
 - Coverage of small and informal enterprises (Area frame surveys)
 - **Data integration, editing and data linking – (4.4)**

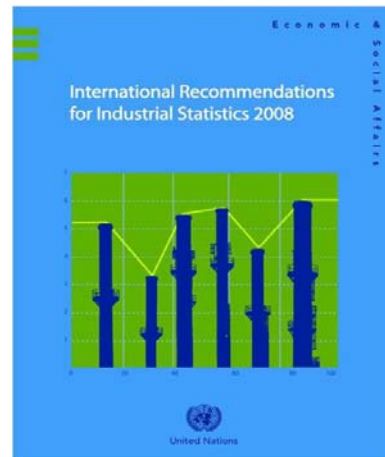
2



Reference

International Recommendations for Industrial Statistics 2008

- Comprehensive methodological framework for the collection and reporting of industrial statistics
- Many recommendations such as guidance on data collection strategy and data compilation practices are generally applicable to development of an integrated economic statistics system for business statistics
- Available:
http://unstats.un.org/unsd/publication/seriesM/seriesm_90e.pdf



3



Diagnostic Framework – survey

Sample surveys as a tool to collect economic statistics

- Technique for obtaining data about a large population of statistical units by selecting and measuring a limited number of units (sample) from that population
 - Conclusions about the total population of units are made on the basis of the estimates obtained from the sample and grossed up with appropriate raising factor.
- Scientific sample designs should be applied in order to reduce the risk of a distorted view of the population
- Sample survey technique is a less costly way of data collection as compared to the census
- It may be used in conjunction with a cut-off point or not

4



Diagnostic Framework - Survey

Household surveys

- Sampling units are households
- Households are selected based on a scientifically designed probability survey (generally multi-stage stratified sampling design)
- Useful variables estimated include
 - Labour force
 - Household production for own final use
 - Income and expenditure of households
 - Household assets and indebtedness

5



Diagnostic Framework – survey

Enterprise Survey

- Sampling frame of enterprises engaged in relevant economic activities – from current up-to-date business register (BR), is a prerequisite
- For countries with no BR, list of enterprises drawn from latest economic census
- Depending upon the source of the sampling frame surveys may also be classified as either *list-based* or *area-based*.
 - In list-based survey the initial sample is selected from a pre-existing list of enterprises (BR),
 - In an area-based survey the initial sampling units are a set of geographical areas. After one or more stages of selection, a sample of areas is identified within which enterprises or households are listed. From this list, the sample is selected and data collected.

6



Enterprise surveys

Information

- A complete list of all economic units – sampling frame
 - mostly establishments for structural data (annual);
 - often enterprises for short-term (monthly or quarterly) production related statistics.
- In the context of data collection for list-based surveys, this list is referred to as a **'list-frame'**.
- But 'list frames' are seldom complete (cut-off).
- Business Register, based on administrative sources and/ or Economic Census

7



Enterprise survey

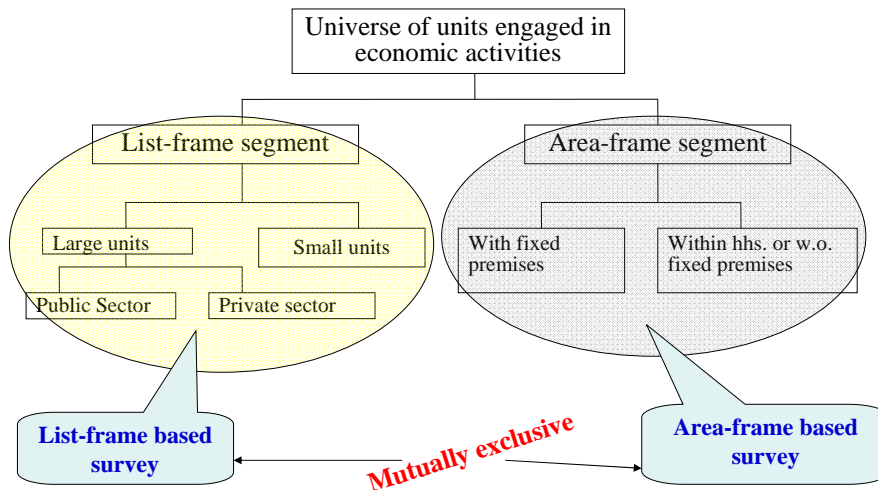
Information:

- ▶ The within-scope units excluded from the list frame are, in some countries, covered using area sampling technique.
- This requires a complete list of well-defined small areas, for example
 - enumeration blocks,
 - enumeration area,
 - village etc.
- This is referred to as an **'area frame'**.
- Area frames are mostly built from the results of Economic / Population Census.
- It provides the data on number of establishments and workers by economic activities (ISIC) for each 'small area'.
- These are used while drawing samples for conducting establishment / economic surveys

8



Enterprise survey – data collection strategies



9



Mixed household-enterprise surveys (1-2 surveys)

Mixed household-enterprise surveys (1-2 survey)

- Sample of households is selected
- Each household is asked whether any of its members own and operate an unincorporated enterprise.
- The list of enterprises thus compiled is used as the basis for selecting the enterprises from which desired data are finally collected.
- Mixed household-enterprise surveys are useful to cover only unincorporated (or household) enterprises which are numerous and cannot be easily registered

10



Data Compilation

- **Data compilation**
 - Comprises more than just aggregating the questionnaire items
 - Statistical offices perform a number of checks, validation and statistical procedures to bring the collected data to the level of the intended statistical output
- Respondents are prone to commit errors while completing a statistical questionnaire
- Data collected through statistical surveys - affected by response and non-response errors of different kinds

11



Data validation and editing

- Integral part of all types of statistical surveys data processing operations
- Required to solve problems of missing, invalid or inconsistent responses
- **Editing**
 - Systematic examination of collected data for the purpose of identifying and eventually modifying the inadmissible, inconsistent and highly questionable or improbable values, according to predetermined rules
 - Essential process for assuring quality of the collected information
- **Types of editing**
 - Micro editing (input editing) - focuses on the editing of an individual record or a questionnaire
 - Macro editing (output editing) – checks are performed on aggregated data

12



Data validation and editing

- Edit checks for detecting errors in data
 - **Routine checks** - test whether all questions have been answered
 - **Validation checks** - test whether answers are permissible
 - **Rational checks** – set of checks based on statistical analysis of respondent data
 - **Plausibility checks** – used to pick up large random errors

13



Non response/Missing data

- **Missing data**
 - Encountered in most surveys
 - Create problems for data editing
- **Types of missing data**
 - Item non-response - data for a particular data item of the questionnaire is missing
 - Unit non-response - selected unit has not returned the filled-in questionnaire

14



Unit non-response

- **Unit non-response may occur due to:**
 - Non existence of the unit included in the survey
 - Lack of appreciation of the importance of the data on part of the respondents
 - Refusal to respond
 - Lack of knowledge how to respond
 - Lack of resources
 - Non-availability of the desired information
- **Ways to minimize unit non-response**
 - Increase the awareness among respondents about importance of surveys
 - Appeal to the respondents to cooperate with the statistical authorities
 - Reminders to the non-respondents and resorting to the enforcement measures laid down in the national legislation

15



How to deal with non-response?

Item non-response

- Ignore all questionnaires with missing values
 - Analysis is confined to the fully completed forms only
 - Not recommended because even the valid data contained in the partially complete forms are discarded
- Imputations
 - Replace one or more erroneous responses or non-responses in a record with plausible and internally consistent values
 - Process of filling gaps and eliminating inconsistencies
 - Means of producing a complete and consistent file containing requisite data
 - Variety of methods

16



How to deal with non-response?

Unit non-response

- Re-weighting - the sample is re-weighted as to include only the responding sample units
- Various forms of imputations – similarly to those used for item non-response
 - Information for the previous periods available for the unit (substitution)
 - Administrative information for the unit

17



Points for consideration

Coping up with changing economic environment – doing more with less resources?

- **Monitor response burden**
- **Need to orient current surveys to capture small and informal sector enterprises adequately**
- **Use of integrated surveys to cover all economic activities rather than separate activity specific surveys – reducing cost**
- **Integration of infra annual surveys with annual surveys**
- **Rotating panel sample design**
- **Need for external technical assistance?**

18



Points for consideration

Data integration, editing and data linking

- Validation and edit checks should be performed on survey data before processing
- Treatment of missing observation and non-response (both item and unit non-response) are done based on sound statistical techniques
- Grossing-up of sample results using raising factor (derived scientifically based on the sample design) for the coverage of the entire in-scope statistical universe
- Rebased series are linked to past series for having a longer time series data on macro-economic aggregates
- Need for external technical assistance?

19



Thank you

20