UN Regional Workshop on Compilation of Basic Economic Statistics, 23-26 July 2007

Economic Large Sample Survey (LSS) Program in South Africa

Presenter: Itani Magwaba
Manager: Economic Large Sample Surveys

Background

- The census of manufacturing was consistently conducted.
- Last economic census was conducted for the manufacturing industry in 1996.
- There was no systematic way of updating the sampling frame.
- The response rate was 43,0%.

Purpose and objectives

Purpose:

- To conduct comprehensive survey program periodically covering the structure, financial performance, inputs and outputs of all industry sectors on a rolling basis, at a detailed industry level.

Objectives:

- To collect financial year data once every three to four years on each industry sector (excluding agriculture, forestry and fisheries, financial intermediation, insurance, government) of the South African economy.
- To publish the first (usually financial) results within one year after the end of the reference year and the final report within twenty four months after the end of the reference year.
- To provide detailed breakdowns on the product composition of 'sales of goods', 'income from services' and 'purchases' than the AFS and short-term economic indicator surveys.

Main users and users

- National accountants use the results to benchmark national accounts aggregates and to compile Supply-Use tables for confronting and reconciling the different approaches to measurement of GDP.
- Price specialists use the data to refresh the composition and weights of the basket of goods and services for producer price indices.
- Industry policy specialists in government agencies use the data to measure the performance and contribution of individual industries to the South African economy, and to evaluate the effectiveness of the relevant policies.
- Individual businesses use the data to analyse their performance relative to the industry as a whole.

Scope and coverage

The program covers **tax registered private and public enterprises** that are **mainly engaged** in the following industries as defined in the Standard Industrial Classification of all Economic activities (**SIC**):

- Mining and quarrying,
- Manufacturing,
- Electricity, gas and water supply,
- Construction,
- Wholesale and retail trade; repair of motor vehicles, motor cycles and personal and household goods; hotels and restaurants,
- Transport, storage and communication,
- Real estate and business services (excludes financial services), and
- Community, social and personal services (except national, provincial and local government activities).

Note: The SIC is based on the 1990 *International Standard Industrial Classification of all Economic Activities* (ISIC) with suitable adaptations for local conditions.

Periodicity and reference period

- The LSS program aims to cover the main industry sectors (manufacturing, trade, real estate and business services) every three years and the other sectors every four years.

- The information would be provided for the financial year of the enterprise which **ends on any date between 1 July 2006 and 30 June 2007, according to the enterprise's usual reporting schedule.**

Example

- 1 October 2006 – 30 September 2007
- 1 January 2006 – 31 December 2006
- 1 February 2006 – 31 January 2007
- 1 March 2006 – 28 February 2007
- 1 April 2006– 31 March 2007

Size groups

The enterprises are divided into four size groups according to the value of their 'business register' turnover. Table A presents the size groups for manufacturing as defined using the Department of Trade (DTI) cut-off points.

Example: Size Group for Manufacturing industry

Large	**:**	**Turnover > R 51 million**
Medium	**:**	**R 13 million < Turnover < R 51 million**
Small	**:**	**R 5 million < Turnover < R 13 million**
Micro	**:**	**Turnover < R 5 million**

Sample Design

- 'Business register' turnover is used as a measure of size for stratification.
- A stratified simple random sampling is drawn.
- The enterprises are first stratified according to the SIC (usually at 4-digit SIC) and then by size of enterprises.
- All large enterprises are completely enumerated.
- Simple random samples are drawn for size groups 2, 3 and 4.

Data items

The following categories of data items are collected:

- **Industrial Classification**
- **Import and export,**
- **Use of ICT,**
- **Details of employment**,
- Trading income,
- Expenditure,
- Profit or loss,
- Inventories,
- Book value of assets,
- **Details of income from services,**
- **Details of sales of goods** (also quantity where applicable), and
- **Details of purchases** (also quantity where applicable).

Questionnaires per industry

Industry	Number
Manufacturing	44
Retail and wholesale	8
Real estate and business services	17
Transport	7

Purpose

- The **main aim** during data collection is to ensure that **quality** data is collected **timely** with a '**high' response rate** (i.e. sufficient quantity).
- Thus, the need to ensure **accurate data** capturing.
- The **editing** of data guarantees that will have quality data.

Organisation

- Data collection and processing is centralised at Head Office.
- Collection is mainly through self-enumeration.
- Personal visits are conducted after request or failure to respond despite repeated telephonic and written reminders. Mostly small and medium sized (in terms of turnover) enterprises are visited.

Organisation (continued...)

- **The workload per enumerator is 300.**

  The total sample size for the 2006 large sample surveys is 12 000 enterprises. The total number of enumerators is 40.

- **The supervisor to staff ration is presently 10.**

  This will be reduced to 6 before starting with the collection of the 2005 large sample surveys.

- **The cost for enumerating one enterprise is R605,00.**

  The budget for the section is R10,3 million and the sample size is 17 000.

Dispatching questionnaires

1. The sample is distributed among the enumerators.
2. The enumerators update and confirm the classification and contact details by telephone.
3. The contact details are updated mainly using phone books, internet, previous surveys and the business register.
4. The questionnaires are then dispatched by mail, fax or internet.

Follow-ups

1. There is a need for continuous follow-ups to remind respondents regarding outstanding questionnaires and to establish if any assistance is needed.

2. Follow-ups are made regularly via telephone and written reminders are sent by fax and/or registered mail.

3. Four written reminders are sent throughout the collection period. The final reminder mentions 'legal action' and is sent by registered mail.

Treatment of various responses

The following are the main response categories:

- Outstanding questionnaire (00).

  A questionnaire can be outstanding because of non-response and no contact.
- Received completed questionnaire (04).
- Received financial statement (05).
- Investigation with contribution (12).
- Investigation with **zero** contribution (13).

Note: Investigation refers to an enterprise which has either 'merged', 'sold', 'split off', 'taken over', 'liquidated during or before survey', 'closed down during or before survey', 'untraceable', 'foreign business' or 'reclassified out of scope'.

Receiving questionnaires and finalising investigations

- Most of the questionnaires and/or investigations are received by fax, followed by mail and e-mail.
- An investigation is only finalised after receiving signed written confirmation from the respondent and approval by the supervisor or survey statistician or manager.
- In some cases, the business register is used to finalise investigation, e.g. an enterprise 'closing down during the survey period'.

Data collection

Preliminary checking

A **'completed' questionnaire** is checked to ensure the following:

- The information received is for the sampled enterprise by comparing company registration number and name.
- The classification is correct by checking main activity. If the classification is incorrect, the enterprise is considered to be 'out of scope'.
- The information is for survey period by checking enterprise's financial year.
- All the relevant questions have been answered.

Capturing System

1. The capturing system is developed using visual basic.
2. The 'editing instructions' are contained within the capturing system.
3. The capturing system is linked to an SQL database.
4. Capturing of questionnaires is done manually through the keyboard.
5. After capturing each questionnaire, a 'warning messages' reported is reflected.

Data processing

Micro-editing

- This process detects errors in data through checking of individual enterprise information.
- Initial micro-editing is done through preliminary checking on receiving completed questionnaire.
- During capturing, mainly range and consistency data edits are done.
- After capturing, ratios are computed and plotted on scatter plot for each stratum (a combination of classification and size group). Outliers are identified and investigated.

Ratios

The following are examples of the computed ratios:

- Total income / 'Register turnover',
- Total expenses / Total income,
- Gross salaries and wages / Number of employees,
- Net profit before tax / Total income i.e. profitability ratio, and
- Sales of services / Sales of goods

Treatment of outliers

- **Large enterprises/ businesses**

  - If a respondent fails to verify outlier, the enterprise is treated as non-response.

  - If the outlier is verified, the information is accepted.

Note: The large enterprises are completely enumerated and thus have a design weight of one.

- **Other enterprises/ businesses**

  - If an outlier (whether verified or not) has an impact on the estimates, it is treated a non-response.

Macro-editing

1. This involves detecting errors in data through the analysis of aggregate data (totals).

2. The aggregated data is compared with data from the following surveys:

2.1 Annual business or household surveys

2.2 Short-term business or household surveys (monthly or quarterly)

Unit imputation
(Compensation for unit non-response)

- **Size group 1 ('Large') enterprises**
  All non-responding large businesses or enterprises are imputed.

- **Size groups 2, 3 and 4**
  The non-responding units are imputed using averages of the strata.

Response rates

2003 Real estate and business services

2004 Construction

2004 Accommodation

2004 Mining

2004 Personal services

2005 Manufacturing

2005 Retail

2005 Wholesale

Dissemination of results

- There are two types of publications:
  - A **'statistical release'** which contains mainly financial information, and/or
  - A **'report'** which contains financial and 'products' information.
- The results are only published at **national** level. No regional figures are published at the moment.

Benefits of the LSS program

- efficient use of resources (staff and financial)

- improved timeliness, reliability and quality of publications, and

- generally satisfied users (especially National Accounts).

Present challenges of the LSS program

- Better co-ordination between LSS and AFS (e.g. on data items).
- Publish regional estimates.
- Improve 'details of sales' on questionnaires, especially for construction and personal services.
- Design more efficient samples.
- Improve response rates.

THANK YOU