Overview of data compilation

Regional Seminar on International Trade Statistics
Bogota, Colombia, 11 – 14 December 2012

United Nations Statistics Division

Karoly Kovacs, Chief, Statistics of International Trade in Services Section

Part III.   Data compilation

Chapters:

12. Introduction and overview of data compilation within the modes of supply framework

13. Integration of data from different sources

14. Compilation of resident/non-resident trade in services statistics

15. Compilation of FATS and the international supply of services

16. Compilation of other indicators for modes of supply

17. Estimation and modeling of missing data, forecasting and back-casting

12. Introduction and overview of data compilation within the modes of supply framework

This chapter reiterates the four modes of supply and the importance of compiling sufficiently detailed resident-nonresident trade in services and foreign affiliates statistics, as well as additional information. It also shows the importance to compile the four modes of supply by distinguishing the modes within each of the resident-nonresident trade in services categories, as well as by distinguishing mode 4 within the statistics of foreign affiliates.

13. Integration of data from different sources

Statistics on the international supply of services are compiled from a number of sources. Therefore, identification of good practices in data integration is of a critical importance. To merge and cross-check data collected from various sources is quite a complex and time-consuming activity. For instance, in order to produce resident-nonresident trade in services statistics, merging ITRS or other administrative data with enterprise or other survey data will require a careful planning.

For FATS, merging information from the structural business statistics and FDI surveys/registers, or other types of linking may need to be carefully tackled. Challenges can even occur while merging different information from the same kind of data source.

The chapter highlights a number of issues that need to be addressed, such as

(a)   matching similar data elements available from different sources;

(b)   working out conceptual differences between sources;

(c)   matching data with different levels of detail;

(d)   working out issues of delays resulting in unsynchronized reception of data;

(e)   overlaps in the information provided and

(f)   the overall efficient use of all data sources.

Besides finding appropriate methodological solutions, some practical solutions may include running training programmes for staff and conducting regular meetings between staff of compiling and source agencies/entities (including staff of large importing and exporting enterprises) to establish stable and efficient working arrangements; and to develop estimation and imputation procedures to deal with the issue of missing data fields

The chapter may consist of the following sections:

- General description of integration of data from different sources

- Consolidation of the data

- Issues encountered when merging data

- Possible approaches and solutions

- Country experiences

14. Compilation of resident/non-resident trade in services statistics

This chapter will recap and extend on the BPM6 compilation guide focusing on the statistical processing aspects. It will provide examples of good practices in compilation of resident/non-resident trade in services statistics.

The chapter consists of several sections:

- General description of compilation of resident/non-resident trade in services statistics

- Compilation for individual service categories (the twelve service categories)

- Modes of supply and resident/non-resident transactions

- Services transactions between related enterprises

15.   Compilation of FATS and the international supply of services

This chapter will describe the compilation of foreign affiliate statistics, including confidentiality issues and possible exchange of data among countries

The chapter consists of several sections:

- General description of compilation of foreign affiliate statistics

- Compiling specific FATS variables focusing on (services products or industries)
  - Sales
  - Output
  - Value added
  - Trade variables
  - Employment (of which intra-corporate movements with breakdown to mode 4, and more generally foreign employment)
  - Other FATS variables (priority and other recommended variables)
  - Breakdown by type of product
  - Compiling services as secondary activities of manufacturing or other industries (i.e. ISIC sections A to F)

Compilation of FATS and the international supply of services

- Dealing with consolidated and unconsolidated data: establishments, enterprise and enterprise groups (in particular in cases of regional headquarters)

- Defining the ultimate controlling institutional unit (UCI) through indirect means (i.e. when information not available in a register or not directly collected)

- Treating the activities of special purpose entities

- Evaluating the validity of reported data

- Confidentiality and suppression of data

- Specific issues for the compilation of inward FATS

- Specific issues for the compilation of outward FATS

16. Compilation of other indicators for modes of supply

This chapter will describe the compilation of the other indicators, in particular non-monetary aspects, of the modes of supply. The compilation of statistics on the number of mode 4 will be discussed in details.

The chapter consists of the following sections:

○ General description of compilation of other statistics on modes of supply (with a particular focus on mode 4)

○ Compiling the number of persons moving under mode 4 regimes

17. Estimation and modeling of missing data, forecasting or back-casting

This chapter describes the use of models and estimates to complement the observed SITS/FATS data which are frequently needed because

(1) *part of the information cannot be collected or can be gathered only at an unsustainable cost;*

(2) *the available sources cannot provide the required coverage, detail, frequency and/or timeliness foreseen by the international standards;*

(3) *the various sources to be combined in the collection system are partially overlapping;*

(4) *limitations of estimating and modeling*

Particularly

(a)   *to impute data gaps and missing replies, and correct suspect outliers etc;*

(b)   *(b) to forecast and compensate for lack of timeliness of data sources;*

(c)   *(c) or to back cast and revise time series backward to incorporate newly available historical data or after a break, for example due to a change in estimation methodology, compilation system, or conceptual standards.*

Models and estimates can be also used to verify the plausibility of the collected data.