COORDINATION OF DATA SETS DISSEMINATED BY DIFFERENT ORGANIZATIONS COVERING THE SAME SUBJECT MATTER

Note by UNSD

1. At its last meeting in Vienna (September 2001) the ACC Subcommittee on Statistical Activities “agreed, that this subject is important and complicated, covering many aspects of how international organizations collect, process and disseminate statistics and try to assure quality”. (…) The Subcommittee “requested the UN Statistics Division to prepare a document on the subject; and invited organizations to provide inputs to the Division within three months, based on a structure that the Division would develop and distribute shortly”. (Report of the ACC Subcommittee on Statistical Activities on its thirty-fifth session, E/CN.3/2002/29, para. 14).

2. The issues to be discussed can be summarized in the following set of questions: (a) What series are known, where there exist discrepancies of data sets disseminated by different organizations covering the same subject matter? (b) How can the differences be explained? (c) What are the value and benefits of identical numbers from different agencies? (d) Is there a tolerable range of difference (one per cent, two per cent …)? (e) Are agencies' methods, definitions and coverage that yield small (or in a few cases large) differences worth perpetuating? (f) Are organizations that are involved in the publication of data series prepared to come to an agreement on the collection/processing and dissemination of data? (g) If not, what are the barriers to agreement (e.g. organizational mandates, history, difficulty of changing series, strongly held professional views on what is "correct", real or imagined political problems) and what can be done to remove these barriers?

3. In order to analyse the reasons for differences in published data on the same subject matter, it is proposed to use the following general structure of the overall data collection/processing and dissemination process:

**Collection**
- national source organizations (can be different)
- definitions
- time of collection

**Processing**
- accept data as provided
- adjust to approximate international recommendations
- estimate for non-reported data
- different conversion factors both monetary and physical measures, different population figures for per capita calculations
- use of data from other national/international sources
- other processing peculiarities to assure quality
- errors

**Dissemination/Presentation**
- different regional and sub-regional aggregates

4. It is suggested that the meeting identify a short list of data series, where problems of discrepancies exist. Known examples are total import and export data, the travel component of the balance of payment or total population estimates. The meeting may wish to ask the organizations involved to analyze jointly the reasons for the discrepancies, using the structure outlined in paragraph 3 above and report back to the group, through UNSD.