Third International Seminar on Early Warning and Business Cycle Indicators

17 – 19 November 2010
Moscow, Russian Federation

Development of Data Hubs for High Frequency Indicators

N.Eagambaram
Central Statistics Office, India

Development of Data Hubs for High Frequency Indicators

N.Eagambaram
Central Statistics Office, India

Introduction

The statistical system is expected to provide the right information to the right people at the right time.   In this endeavour communication and dissemination of statistics play an important role and it is more so in the present context of introducing highly frequent economic indicators. In this paper we briefly describe the statistical system in existence specifically with reference to India and suggest some possible ways for setting up a data hub for disseminating High Frequency Indicators (HFI).

2.Statistical System in existence

In most of the countries with a federal set up, statistics is a common subject matter of business for both the central and regional governments. There are a number of ministries at the centre each dealing with a particular subject.   Statistical units exist in individual ministries catering to the administrative needs according to the mandate of the ministry. The regional governments have similar set up to take care of the regional requirement. The National Statistical Office (NSO) at the centre is entrusted with the responsibility of coordination, prescribing standards, compilation of national level statistics and providing guidance to the central and regional agencies in the matter of production and dissemination of statistics of importance at the national and regional level.   Periodical surveys and censuses are also conducted by NSO and other agencies under the guidance of NSO to generate time series data on important topics for planning and policy making.

2.1. Indian System

In India too, statistics is a common subject matter for   both the Union and State Governments. The Central Statistics Office (CSO) in the Ministry of Statistics & Programme Implementation (MOSPI)   is the nodal agency at the national level for prescribing standards, coordinating statistical work for uniformity and consistency, interacting with international agencies and disseminating core statistics at the national level. National Sample Survey Organisation (NSSO) who conduct periodic nation-wide surveys on various socio-economic subject matters and Computer Centre (CC) engaged in data processing and maintenance of data warehouse and website are the other wings of MOSPI. Above all the National Statistical Commission is the highest national level agency which oversees the overall development of the statistical system which comprises the systems in the Union Ministries including CSO and the State and Union Territory Governments.

The State Statistical Bureaus (SSB) in the States and Union Territories are the nodal agencies engaged in similar activities like CSO, at the State level. At the national level, the individual

ministries have statistical offices looking after the specific requirements of the concerned ministries. Data dissemination is done by CSO and other agencies through periodical publications.

The statistical system and the data flow among the agencies are diagrammatically presented below:

NSSO: National Sample Survey Organisation
CSO : Central Statistics Office
CC   : Computer Centre
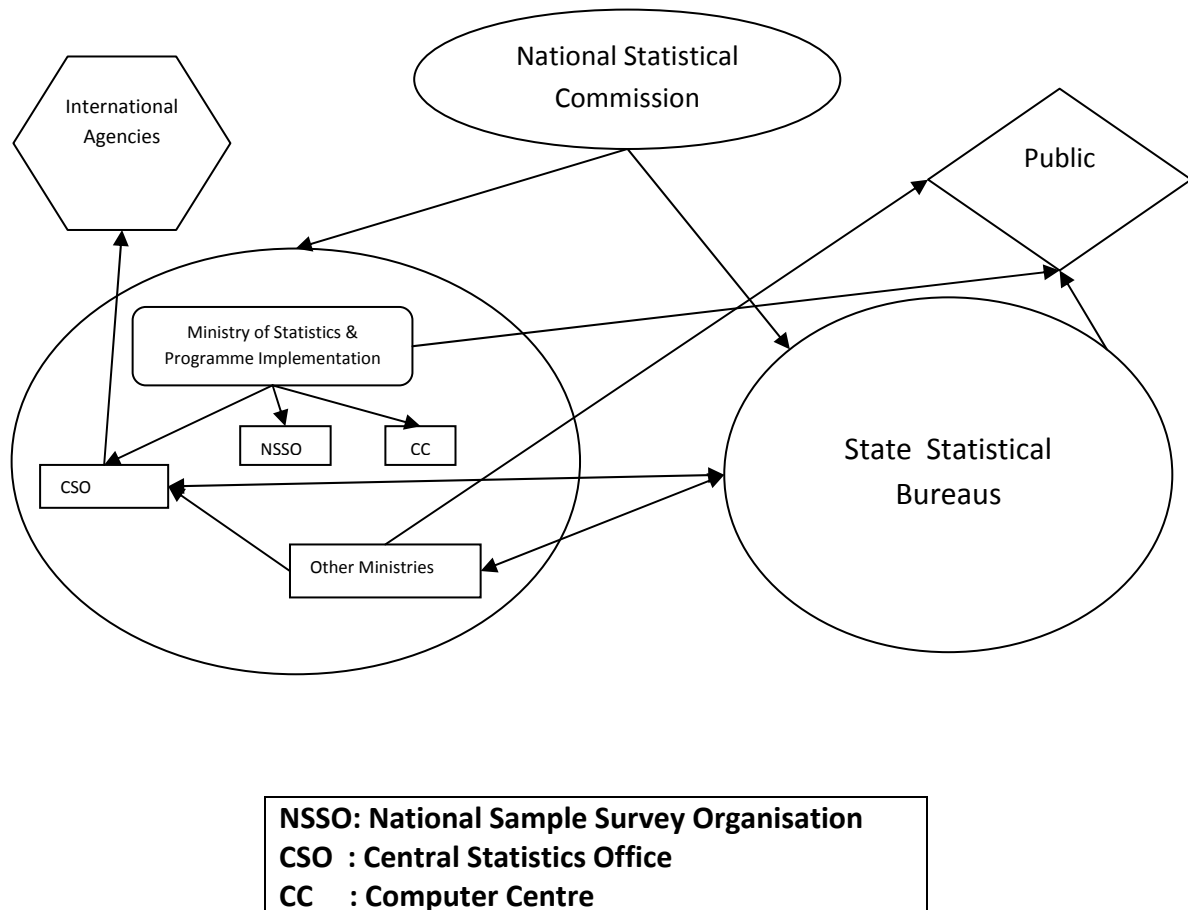
The central ministries and SSBs maintain their data on computers. Many of these agencies are equipped with client-server systems. But a computer network for online data exchange between the data agencies is yet to be established.

3. Issues on Data Hub for HFI

Problems to be addressed in the establishment of a data hub and making it operational are coordination, data collection and data exchange, identification of a proper mode of dissemination, constant monitoring and guidance, and above all commitment to compile and publish the indicators as per schedule.

3.1 Coordination

Coordination by the nodal agency is an important item of work in compilation and dissemination of HFIs because there is no single agency who compiles all the HFIs. Even in a well developed statistical system coordination is not a simple task as data are to be collected from within and outside the system. Since data are received mainly in the conventional paper mode for reasons of authenticity and responsibility, it adds to the work pressure of the coordinator to collect the frequent data sets on HFIs.   Lack of commitment in the new venture on the part of data provider is likely to create delay in response. A nodal agency is to be identified and a group for implementation of HFIs is to be set up to prepare the ground work and obtain agreement and commitment from the data agencies at the country level.

3.2 Collection and exchange of data

 Some of the HFIs require frequent collection of direct field data. However, use of electronic gadgets in censuses and surveys are only in the experimental stage in many countries. Data processing is done using computers but data dissemination is still done mainly in the conventional mode of printed publication. Reports are published in the web sites at the most as static pages. Even within the statistical system data exchange takes place among the agencies through ordinary mail. Electronic mode is at the most restricted to emails. Online data transfer and web based applications have not yet taken roots. Data to international agencies are also provided by conventional mail. It is seen that the international agencies also collect data from countries through conventional mail. For the purpose of HFI, it is essential to use computers and computer network to collect, check and process the data for compilation. Where such infrastructure is not available, necessary guidance and assistance from the international agency is required for establishing the system.

4. Suggestions on Data Hub for High Frequency Indicators

Arrangement for collection and compilation of additional data for HFIs by different agencies enforcing uniformity of standards and procedures, preparation of simple and composite indicators and disseminating them for use within the country and by international agencies have to take place in the back drop of constraints and limitations of the existing statistical system.

4.1 Local Data Hub

The data hub has to cater to the local needs of the country as a first priority and provide data periodically to the international and regional agencies according to the prescribed time schedule. Though the data generation agencies for HFIs are scattered over different organisations and regions, the major part of these data and indicators are either compiled directly or collected and maintained by the NSO. Therefore, the local database for HFIs can

be maintained by NSO by coordinating with all concerned within the country. The mode of collection of data by NSO can vary depending on the data communication facility available with the source agency.

 Maintenance of HFI data hub by NSO involves:

(i) Consultation with the source agencies and getting their agreement for cooperation in establishing the system for HFI

(ii) Prescription of methodology and procedure for collection of data and compilation of indicators

(iii) Prescription of the format for reporting

(iv) Assisting the source agencies on creation of IT infrastructure for speedier transfer of data

(v) Create a web page for publishing the data.

(vi) Compilation composite indicators, if any, and arrange to upload the data periodically as per schedule.

(vii)   Establish connectivity with the concerned international agencies for supply of data

Just as UNSC is arranging for collection and dissemination of data at the international level, at the national level also a permanent group should be set up to take care of the local data hub. Each data source should be bound by commitments to provide the data in time.

The format, periodicity and standards are to be finalised in consultation with the participating countries which may be coordinated by UNSD.

4.2 Data for international agencies

International organisations such as UNSD and FAO collect data from countries in only in physical mode. Data on National Accounts, Crop and livestock products are collected in standard formats. Sometimes the international agencies approach the individual agencies directly for data without routing the request through NSO. While annual statistics can still be collected in this manner, HFIs require a faster mode of transmission for quicker compilation and dissemination. A system that requires minimum manual intervention is to be put in place for this purpose.

4.3 SDDS as an example

The Special Data Dissemination Standards (SDDS) established by IMF which includes several of the HFIs in its coverage, has been a successful venture that has stood the test of time. Metadata and time schedule have been worked out in consultation with the participating

countries and commitment obtained from the participants to supply the data of prescribed standard in time.

India subscribed to SDDS in 1996. Data on Real Sector, Fiscal Sector, Financial Sector, External Sector and population are put up in the websites of the respective Ministries. Advance Release Calendars (ARC) and National Summary Data Pages (NSDP) are also displayed in the respective websites. The nodal agency for hosting the NSDP for all these sectors is the Ministry of Finance vide. their web page, 'www.finmin.nic.in/stats_data/nsdp_sdds/index.html' which is the page linked to the IMF website. The standards and quality of data are made available in Data Standards Bulletin Board (DSBB) of IMF which are also updated as and when needed. Contact persons for the country NSDP, ARC and detailed data are also identified. The arrangement for data dissemination is merely on the basis of linking of the relevant web pages.

The example of SDDS can be copied for HFI purposes too by taking NSO as the nodal agency. This seems to be a solution which can be implemented immediately.

4.4 Regional Data Exchange System(RDES) of FAO

Under the Asia and Pacific Cooperation on Agricultural Statistics (APCAS), Regional Data Exchange System (RDES) has been set up. The web site of RDES contain the regional agricultural data along with meta data. There is a provision for online data entry by the source agency relating to sub-regional data. The data are collected from the source agencies called focal points in spreadsheet format and loaded onto the RDES database. This system has the advantage for further processing and tabulating the data for inter-country or inter-regional comparison. The database built up over time would also serve analytical needs of researchers and other users.

The RDES-type system can be followed at the national level also for establishing the local data hub. However, the data source agencies should be trained and motivated to feed the indicator data periodically in time adhering to the time schedule. Constant monitoring and timely intervention are to be done by the national agency maintaining the data hub.

4.5 An ideal way

Information Technology has penetrated all fields of human activity. Ideally, a separate database and web server may be maintained by UNSD for collecting national data. Connectivity may be provided to the national data hubs for smooth transfer of data into the UNSD database. A query system may be developed for data retrieval. Inter-country comparison and derivation of composite or analytical indicators will become possible in this set up. The database will become a storehouse of time series data over time. The system can serve as a medium of communication in all matters relating to HFI including guidance on various aspects compilation of HFIs. The HFI system will have to be maintained by an expert team in UNSD for making improvements and bringing about better comparability of indicators across regions over time. A similar system can be thought of at the national level for the local data hub.

5. Conclusion

1. Coordination with source agency and timely collection of data especially in the context of HFI is real task that would have to be faced by the local nodal agencies

2. A statistical system with proper arrangement for data flow within a country is a pre-requisite for implementing data hub.

3. Local group in the country is to be set up for implementing and monitoring the generation and dissemination of HFI in accordance with the standards.

4. The SDDS model based on linking the static pages is an immediate solution as this model is being successfully implemented by the IMF. NSO can be identified as the country nodal agency who would arrange to publish the data in the NSO web site.

5. The model such as RDES is also a feasible solution provided there is commitment on the part of data providers and regular monitoring by the international agency. Here too, NSO can be the country nodal agency with the responsibility to publish the data in the NSO web site and feed the same into the international server.

6. An ideal mechanism is to have a database at UNSD and provide web connectivity to the local databases for transferring data online in which case a similar structure at the country level also is required to be established.

7. There should be a permanent arrangement for providing guidance and assistance to countries in ensuring the minimum level of uniformity and comparability in the indicators across countries and in the construction of those indicators which are not being compiled in a country. Constant monitoring of the system by the international agency is required to make the system successful.