Redesigning India's ailing data system

Writer – R.B. Barman (former Chairman, National Statistical Commission)

This article is related to General Studies-Paper-III (Indian Economy)

The Hindu

21 Jan., 2020

The present national accounting and analytical framework misses out on many key dimensions of a complex economy.

The new series of Gross Domestic Product (GDP) figures with 2011-12 as base, released in 2015, has not gone well with analysts; the withholding of employment-unemployment data for some time and consumer expenditure data, which is not released, added to this unease. Bringing the National Sample Survey Office (NSSO) under the fold of National Statistics Office, altering its long-standing arrangement under the Governing Council and then National Statistical Commission, triggered suspicion. As official statistics is a public good, giving information about the state of the economy and success of governance, it needs to be independent to be impartial.

Wide-ranging impact

GDP covers all productive activity for producing goods and services, without duplication. In effect it adds apples and oranges, tractors and sickles, trade, transport, storage and communication, real estate, banking and government services through the mechanism of value. The System of National Accounting (SNA) is designed to measure production, consumption, and accumulation of income and wealth for assessing the performance of the economy. GDP data influence markets, signalling investment sentiments, flow of funds and balance of payments. The input-output relations impact productivity and allocation of resources; demand and supply influences prices, exchange rates, wage rates, employment and standard of living, affecting all walks of life.

The data on GDP are initially estimated at current price and then deflated for constant price for comparability of data over time. It is necessary to separate out price effect to adjust value for real volume for comparison over time and sectors. There is a way of adjusting price effect through appropriate price index. The present series encountered serious problems for price adjustment, specifically for the services sector contributing about 60% of GDP, in the absence of appropriate price indices for most service sectors. The deflators used in the new series could not effectively separate out price effect from the current value to arrive at a real volume estimate at constant price. Price indices going into a low and negative zone in 2014-17 distorted real growth.

The shift from establishment to enterprise approach, replacing Annual Survey of Industries (ASI) with Ministry of Corporate Affairs MCA21 posed serious data and methodological issues. The use of MCA21 data and blow up factors thereof without weeding out defunct enterprises, and then insufficient work on mapping of comparable ASI data, followed by similar survey on services sector enterprises were another major lacunae.

Unchanged approach

The approach for collection of data remains largely the same for long: price and production indices are constructed using a fixed base Laspeyres Index, yield rate for paddy is estimated by crop cutting experiments, and the organisation of field surveys for collection of data on employment-unemployment, consumer expenditure, industrial output,

assets and liabilities continue. When productivity and remunerative price of output are major concerns for agriculture, it is necessary to collect data on factors such as soil conditions, moisture, temperature, water and fertilizer use determining yield, impact of intermediary and forward trade on farm gate price and so on. For example, Israel collects these data for analysis to support productivity. The initiative under e-governance enabled the capturing of huge data, which need to be collated for their meaningful use for production of official statistics. The process for collection and collation of data needs modernisation using technology.

Data logistics

Along with GDP, we need data to assess competitiveness, inclusive growth, fourth-generation Industrial Revolution riding on the Internet of things, biotechnology, robotics-influencing employment and productivity, environmental protection, sustainable development and social welfare. Hence GDP data needs to be linked with a host of other data for deeper insight. We need to re-engineer the existing system, creating an integrated system populated with granular data. The country is vast, heterogeneous. There are non-linearities and path dependence, which should be considered while setting goals for development, reducing regional imbalance. To pursue the goal of a $5-trillion economy by 2024-25, harnessing demographic dividend, we must tap underused resources for demand creating investment, which require data to pursue policy right from a district and evaluate performance for efficiency including governance.

We cannot reconcile data inconsistencies by setting up committees alone. We need systems which have the capability to sift through a huge volume of data seamlessly to look for reliability, validity, consistency and coherence which, I am afraid, will be difficult without the aid of a versatile data warehouse as a component of big-data technology. Also such committees should have the support of a dedicated team for audit, and the ability to implement decisions by cutting red tape. This is what has been wanting as thoughtful and well-meaning key recommendations of the Rangarajan Commission and subsequent recommendations from 2006 onwards by successive National Statistical Commissions, faced stumbling blocks in implementation. What is the guarantee that the recently constituted committee will succeed in its effort to restore credibility? It is as much a system and has implementation capability as the expertise behind it.

The present and future

The present national accounting and analytical framework misses out on many important dimensions of the economy and its complex character as an evolving economy that is constantly experiencing technological and institutional transitions and power plays in a market economy. We need a new framework for analysis for such a complex system and evolutionary process. There is a question of growing market power, automation, robotisation and other labour-replacing technologies affecting profitability, structural change and general welfare.

We need to find alternative avenues for the unemployed and jobs lost. In order to inject efficiency and stability, we need to have detailed data on how: markets clear, prices are formed, risks build up, institutions function and, in turn, influence the lifestyle of various sections of the people. We also need to know in greater detail about market microstructure and optimality therein, the role of technology and advanced research, changing demand on human skills, and enterprise and organising ability, which are all complex. The growing inequality and concentration of wealth in a few hands to the detriment of social welfare needs to be arrested at the earliest. The deadweight loss caused to the economy through monopoly power, inefficient input-output mix, dumping, obsolete technology and production mix must be contained.

The consensus macroeconomic framework of analysis assumes symmetric income distribution, and does not get into the depth of structural issues, as it focuses on a trend-cycle decomposition of GDP for growth and stability in market parlance and a trickledown effect for percolation of income. This framework is questioned by many. The alternative to be realistic for the real world must rest on two pillars: the micro-behaviour of individuals, and the structure of their mutual interactions. In the changed situation of availability of micro data, we need to build a system to integrate the micro with the macro, maintaining distributional characteristics.

Data is the new oil in the modern networked economy in pursuit of socio-economic development. The economics now is deeply rooted in data, measuring and impacting competitiveness, risks, opportunities and social welfare in an integrated manner, going much beyond macroeconomics. We have a commitment to produce these statistics transparently, following internationally accepted standards, tailor-made to suit local conditions, for multi-disciplinary analytics. As these statistics reflect on the performance of the government, it is necessary that its independence is maintained scrupulously.

Expected Questions (Prelims Exams)

1. Consider the following statements:

1. MCA-21 is an e-Governance initiative of the Ministry of Corporate Affairs, Government of India.
2. The performance of the economy is measured through the System of National Accounts (SNA).
3. GDP data affects the market, flow of investment and balance of payments.

Which of the above statements are correct?

(a) 1 and 2 (b) 2 and 3

(c) 1 and 3 (d) 1, 2 and 3

Note: Answer of Prelims Expected Question given on 20 Jan., is 1 (d)

Expected Questions (Mains Exams)

Q. 'In the current world order, any country can move on the path of development only with the logical use of data.' In the context of this statement, discuss India's data system and related institutional structure. (250 words)

Note: - The question of the main examination given for practice is designed keeping in mind the upcoming UPSC main examination. Therefore, to get an answer to this question, you can take the help of this source as well as other sources related to this topic.