

*Inaugural Address at the 17th Statistics Day Conference**

Shaktikanta Das

I am happy to participate in the 17th annual 'Statistics Day Conference' of the Reserve Bank, and also to launch the 'Centralised Information Management System' or CIMS, which is our next generation data warehouse. Over two decades ago, the Reserve Bank was among the pioneer central banks to set up its data warehouse. Considering the developments during the interregnum, it is natural that we migrate to a new platform with a more enriched orientation.

Celebration of the National Statistics Day, which marks the birth anniversary of Professor Prasanta Chandra Mahalanobis, provides an opportunity for sensitising the general public, especially the younger minds, about the discipline of statistics, which provides scientific basis for taking informed decisions across diverse domains. Prof. Mahalanobis was instrumental in institutionalising statistics in India, both as an academic discipline and as a policy devising tool. He always approached the subject as an applied science, capable of addressing real life questions. We are reminded of practical application of statistics, such as, his analysis of half a century data on floods, which influenced the construction of the Hirakud dam, and the second Five-year Plan model which focussed on rapid industrialisation of the Indian economy. All these demonstrated how statistics can address complex problems and drive progress. As we pay our tribute to Prof. Mahalanobis on his birth anniversary, we continue to draw inspiration from the sagacious endeavours of this great visionary.

* Inaugural Address by Shri Shaktikanta Das, Governor at the 17th Statistics Day Conference organised by the Department of Statistics and Information Management, Reserve Bank of India on June 30, 2023, Mumbai.

I would also like to take this opportunity to congratulate Dr. C.R. Rao, the living legend in statistics and a close associate of Prof. Mahalanobis, who has been selected for the prestigious and long overdue International Prize in Statistics in 2023. Incidentally, we have two of his distinguished students, Professor S.R.S. Varadhan and Professor Rajeeva Karandikar amidst us today. They will be talking about interesting areas in statistical theory and applications, later during the day.

Evidence and analysis are core inputs in policy formulation. Of late, macroeconomic policy making and monitoring processes have become more data intensive, with considerable reliance being placed on detailed study of developments, interlinkages among factors, identification of patterns, forecasts of possible path and scenario analysis – all aided by the discipline of statistics. A prerequisite of such analysis is the availability of timely and credible data with the three Cs of data quality, *i.e.*, completeness, correctness and consistency.

Unlike the economic variables that are directly compiled from transaction systems, which are timely and robust, certain core macroeconomic aggregates - such as, gross domestic product (GDP) growth and price inflation, where compilation process is dependent on multiple channels - are available with a time lag globally. Also, their early estimates, which are indeed very useful, are compiled with limited set of inputs and partial data and are prone to multiple and sometimes significant revisions. Policymakers who use them as inputs, however, do not have the luxury of revising decisions in retrospect¹. Monetary policymakers supplement official estimates with information on auxiliary variables to have firmer

¹ Athanasios Orphanides (2001). "Monetary Policy Rules Based on Real-Time Data". *The American Economic Review*, Vol. 91, No. 4 (Sep., 2001), pp. 964-985.

assessment and minimise policy errors emanating from data revision. Statistical information is also used by businesses and households in making assessment of the economic situation and firming up their near-term expectations.

To our advantage, technological developments have kept pace with the rising dimensions and depth of economic activities to support their detailed monitoring. Advances in remote sensing, automation, digitisation, information management ecosystem, text mining, natural language processing, artificial intelligence and nowcasting provide us with quick and comprehensive information on activities. Their optimal utilisation has supported and can further refine compilation of macroeconomic aggregates and increase their efficacy in navigating through uncertainties.

The Reserve Bank uses statistical methods in almost all its core functions and is both a compiler and a user of macro-financial statistics as well as other economic data collected through regular surveys. The Reserve Bank follows latest global prescriptions and best practices, and pursues standardisation across domains to generate consistent, comparable and harmonised statistics. We treat data as 'public good' and are disseminating increasingly more data in public domain for use by analysts, researchers and general public. Our preference is for general dissemination over meeting individual requirements.

The Reserve Bank established its first enterprise-wide data warehouse – the Central Database Management System (CDBMS) – which was accessible to its internal users since 2002. A large part of this data system was placed in the public domain as the 'Database on Indian Economy (DBIE)' portal in November 2004. Over the years, DBIE has evolved from a simple data repository to an information processing and management system, which has become the Reserve Bank's data dissemination

platform. The DBIE remains very popular among domestic and international researchers, analysts, and general public, especially students. It received over 2.5 lakh hits in May 2023.

The Reserve Bank's Regulations Review Authority 2.0 (RRA 2.0) has recently made several recommendations on streamlining reporting mechanism and reduction in regulatory compliance burden. Many of these recommendations have already been implemented and others are in various stages of implementation. A major recommendation on system-based submission of the remaining email-based reporting will be implemented through the Centralised Information Management System (CIMS) in the coming months.

Our investment in technology for information management, periodic reviews, continuous engagement with reporting entities and technological upgradation at their end, have paid rich dividends in terms of improving coverage, quality, and timeliness of data. During the COVID-19 lockdown period, our reporting system ensured business continuity; the flow of validated information was seamless; the 'work from home (WFH)' environment was actively supported; and the public dissemination of information went uninterrupted.

With today's launch of CIMS, we embark on a major change in our information management framework for handling the massive data flow, aggregation, analysis, public dissemination and data governance. This system uses state-of-the-art technology to manage Big data and will serve as a platform for power users to carry out data mining, text mining, visual analytics and advanced statistical analysis connecting data from multiple domains, such as, financial, external, fiscal, corporate and real sectors as well as prices. In short to medium term, it would lead to a paradigm shift in the Reserve Bank's economic analysis as well as supervision, monitoring and enforcement across multiple domains.

The new system is starting with reporting by scheduled commercial banks and will be gradually extended to urban cooperative banks (UCBs) and non-banking financial companies (NBFCs). Incidentally, with the CIMS going live, the first weekly statistical supplement (WSS), which is the Reserve Bank's weekly data release on its own operations and on developments in banking and financial markets, was compiled and processed in the CIMS for the week ended June 23, 2023. It will disseminate more data for public use and will also support on-line statistical analysis by external users at their end. Regulated entities will also have access to their past data and their assessment on quality parameters in the new system.

Any transition in system with multiple dimensions involving large number of entities is prone to teething trouble and therefore, our teams will support the reporting entities for smooth transition, wherever necessary. Several new features will also be augmented in the coming months.

This conference is being organised against the backdrop of two international events with focus on

statistics; one, the principle of data for development is an integral part of the work stream under India's on-going presidency of G20; and second, the upcoming membership of India in the United Nations Statistical Commission (UNSC) after a gap of two decades. I note that results of four research papers covering various aspects of G20 economies will also be presented today.

Let me now conclude with the words of Prof. Mahalanobis: "*We shall naturally devote closer attention to the collection and analysis of data relating to India, but we shall try to study all Indian questions in relation to world problems*"². Drawing inspiration from these words, let me also give a message to our statisticians in the Reserve Bank: as you look for a wider canvas in your profession, I urge all of you to abide by the spirit of these words of Prof. Mahalanobis.

I wish all success to today's deliberations, which I am sure, will ignite our teams, particularly the younger officers, to strive for professional excellence in their commitment to serve the nation.

Thank you.

² Editorial of the first issue of *Sankhya: The Indian Journal of Statistics* (1933).