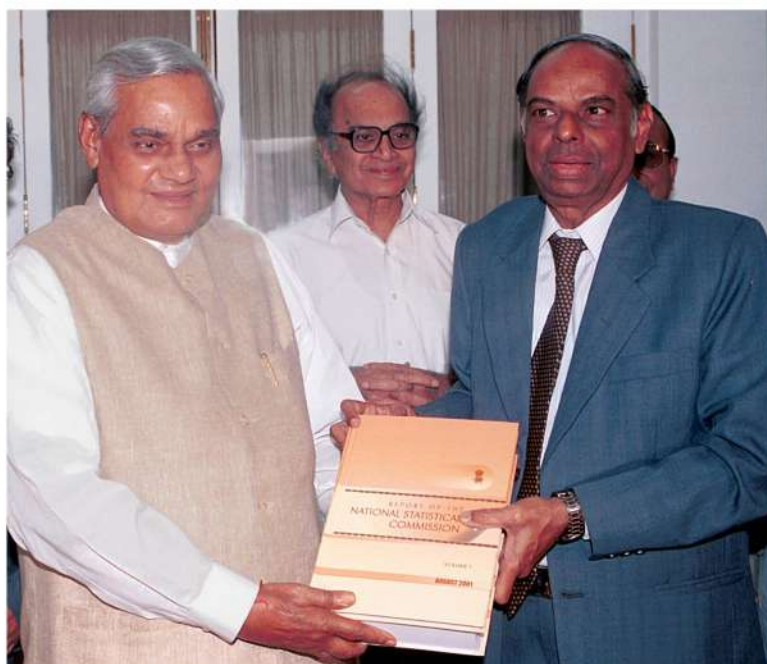


Policy Watch

No. 16



Credible Data for the Public Good: Constraints, Challenges, and the Way Ahead

P. C. Mohanan

Former Acting Chairman
National Statistical Commission (NSC)



THE HINDU CENTRE

for

Politics and Public Policy

The Hindu Centre for Politics and Public Policy, Chennai, is an independent platform for exploration of ideas and public policies. As a public policy resource, our aim is to help the public increase its awareness of its political, social, and moral choices. The Hindu Centre believes that informed citizens can exercise their democratic rights better.

In accordance with this mission, The Hindu Centre's publications are intended to explain and highlight issues and themes that are the subject of public debate and aid the public in making informed judgments on issues of public importance.

Photo Caption: Dr. C. Rangarajan (right), the Chairman of the National Statistical Commission, presenting the Report of the Commission to Prime Minister Atal Bihari Vajpayee, on September 5, 2001. The Commission had recommended far-reaching restructuring of the country's statistical system, which remain elusive. **File Photo:** Kamal Narang / The Hindu BusinessLine

© The Hindu Centre for Politics & Public Policy, 2022

All rights reserved.

No part of this publication may be reproduced in any form
without the written permission of the publisher.

**Credible Data for the Public Good:
Constraints, Challenges, and the Way Ahead**

P. C. Mohanan
Former Acting Chairman
National Statistical Commission (NSC)



THE HINDU CENTRE

for

Politics and Public Policy

TABLE OF CONTENTS

I	Introduction	1
II	The Developing Framework of Official Data	3
III	Evolution of Official Statistics in India	7
IV	Growing Concern over Data Inadequacies in Independent India	11
V	Reviving Interest in Official Data	17
VI	Data as a Political Tool	20
VII	The Federal Disconnect	24
VIII	Institutional Inadequacies: Causes and Implications	27
IX	Way Forward	31

ABSTRACT

Bits and bytes of information propel today's knowledge society. This data revolution is as transformational as it is multi-dimensional. India, however, remains a laggard and is yet to harness the full potential of data for the public good.

In this Policy Watch, **P. C. Mohanan, former Acting Chairman, National Statistical Commission (NSC)**, takes the reader through the data collection, analysis and dissemination process in India. In particular, he points out the deficiencies in the institutional, implementational, and procedural elements of the country's official statistics machinery. For a country endowed with a multiplicity of resources that are matched by the problems that confront it, the scientific use of data to address peoples' issues has often been subverted for either political reasons or because of the inability of the structures that are in place to deliver timely and credible data for decision-makers.

As the rest of the world races ahead by adapting newer technologies and creating independent bodies that ensure credibility of data, India appears to not only stagnate, but regress as well. The way out, Mohanan says, is to harness the available technologies in a meaningful manner, improve statistical literacy, and insulate the statistical system from political vested interests.

Keywords: National Statistical Commission, Census of India, Rangarajan Committee, Statistical Literacy, Ministry of Statistics and Programme Implementation, India Statistics, Data Governance Policy

I. INTRODUCTION

“Statistics is a gateway to knowledge and the progress a country makes is vitally dependent on the efficiency of its statistical system.”

C. Radhakrishna Rao

(Statistics, Statisticians and Public Policy Making, 1983)¹

Any observer of independent India’s development process cannot ignore the role played by statistics in the country’s development process. The centrality of empiricism in planning and policy making has been highlighted by many eminent international commentators who were privy to these developments and are now being rediscovered through careful research using archival materials². What is striking in these narratives is the close interlinking of statistical theory and practice and the official patronage it received, leading to the establishment of outstanding institutions and mechanisms for generation, analysis, and dissemination of official data.

Seventy five years of independence, now marked by national celebrations, is also an appropriate timescale to critically assess the state of India’s official statistics. The digital revolution and the all-pervading use of ‘data science’ substantially impact discussions comparing the past with the present. This vastly expanded scope of data-centric discussions, therefore, needs to be somewhat pruned to highlight specific issues plaguing official data systems. This Policy Watch aims to critically examine the current state of official statistics and the mechanism that produces it. The developments in Statistics as a science, where the contribution of Indian Statisticians continue to be of a very high order, is beyond the scope of this article.

This Policy Watch first takes note of the key developments relating to data over the past decade which saw transformative advances in Information Technology. It then briefly recalls the evolution of the statistical system before and after independence especially in the light of the political structure that India gave itself through its constitution. The changing role of the Union and the States becomes critical in this political framework, which has an impact on the administrative agencies involved in data collection. It, thereafter, outlines the slow decline in the outcomes that were expected from the initial developments

¹ Rao, C.R. 1983. *Statistics, Statisticians and Policy Making*, Sankhya: The Indian Journal of Statistics, Series B (1960-2002), Indian Statistical Institute, August, Vol 45. No 2, pp. 151-159. [<https://www.jstor.org/stable/25052287>].

² Dandavate, M. 1999. *Statistics, Planning and Development*, Sankhya: The Indian Journal of Statistics, Series B (1960-2002), Indian Statistical Institute, August, Vol 61, No. 2, pp. 229-236. [<https://www.jstor.org/stable/25053080>].

and the failure of the statistical system to meet the growing demand for data from within and outside the government. The next two chapters provide the current state of affairs, including some recent debates on data, and some prescriptions.

Finally, another important context is the sudden interest in world of data, in particular the sharing of data collected both as part of official statistics and through the nearly ubiquitous digital platforms. While the former is still confined to traditional modes of collection, the latter are largely 'harvested' from individuals when they use digital platforms to purchase and pay for goods and services. In addition, the digital footprints left by visitors to the internet constitute a large body of data that are already used by private sector to increase their markets and capture new ones. Against this backdrop, the Policy Watch also looks at the proposed Draft Data Governance Framework.

II. THE DEVELOPING FRAMEWORK OF OFFICIAL DATA

“Indian Official Statistical System is wrapped in the cobweb of time, which needs thorough revamping. The silos as legacy systems of the past are not only not efficient but have often created hurdles on absorption of technology to modernize the systems and processes which are needed to maintain high quality, consistency, coherence and timeliness of collected data. It is necessary to support decision making at all levels of governance and also inculcate a collaborative approach giving a better problem solving ability on credible and responsive public policy. The stake holders, that is, enterprises, people in general, civil societies and international institutions also need credible data as public good. ...[T]here is a pressing need for ... strengthening of state statistical systems”.

National Statistical Commission

Annual Report, 2017-18

(p. 57-58, para. 6.10)

When the world of knowledge is in the thick of a data revolution, it would only be appropriate to first present an update on “what is happening” in the use of statistics for social and economic development. The awareness that official statistics has a role that goes beyond public administration gained currency with the latest phase of globalisation. Technological advancements in the field of information processing and dissemination accelerated the demand for data from beyond ‘official’ boundaries, creating newer dynamics between data owners, creators, and users.

The UN General Assembly, on January 29, 2014, endorsed the Fundamental Principles of Official Statistics which were adopted by the UN Economic and Social Council on July 24, 2013. The first of these 10 principles, which defines the relationships between data, society, and official agencies, states:

“statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens’ entitlement to public information”.³

³ **United Nations. 2014.** *Fundamental Principles of National Official Statistics*, United Nations Statistics Division, January 29. [<https://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf>].

National statistical agencies in most countries have adopted the UN's Fundamental Principles of official statistics as the guiding principles in articulating their mission. In 2016, the Government of India also adopted these principles. It claimed that doing so will bring professional independence, impartiality, accountability, and transparency in collection, compilation, and dissemination of official statistics in accordance with international standards. This, however, has been easier said than implemented, owing to a multiplicity of factors.

The World Development Report (WDR)⁴ 2021, titled *Data for Better Lives*⁵, highlights the importance of the Enlightenment era of the 18th century for a major transition in the use of data: from using it merely to collect revenue, organise the military, and monitor employment to further illuminate political and popular understanding of societies. It goes on to provide examples of the use of data to understand the spread of poverty (by the British Sociologist, Rowntree) and the incidence of diseases (the Yellow Fever outbreak in New York city at the end of the 18th century, and the 1854 cholera outbreak in London).

More importantly, WDR 2021 also makes the difference between data collected for “commercial purposes”, which it calls “private intent data” and those for “public purposes, called public intent data”.⁶ This distinction gains importance with the widespread use of digital technologies and the reach of social media for various transactions. The collection of data harvested through such new modes has opened up many interesting prospects and raised important questions. One such issue concerns the control of personal data. This explosion in the use of personal data for planning, decision making, and increasing markets is propelled by the rapid expansion in technologies in mobile telephony, data transmission, and the usage of spectrum (owned by sovereign states) for telecommunication services (provided by a few entities) to harness data from a large base of (millions of) individual subscribers.

The once-in-a-century COVID-19 pandemic, which led to lockdowns of varying degrees, also saw the increased utilisation of telecommunications to gather personal data through smartphones or by tracking the location of individuals using call detail records (CDRs). Although data protection laws date back to the 1970s⁷, the pandemic, which touched every part of the world, its inhabitants and governments, further

⁴ One of the major global landmarks in the mainstreaming of data for the study of economic growth and development was launch of the annual World Development Reports (WDR) by the International Bank for Reconstruction and Development (the World Bank) in 1978. Every WDR is specific to a theme, and contains a large body of information, including data – both general and specific to the theme of that particular year.

⁵ **International Bank for Reconstruction and Development / The World Bank. 2021.** [World Development Report 2021 – Data for Better Lives](https://www.worldbank.org/en/publication/wdr2021), World Bank Group, Washington DC.[https://www.worldbank.org/en/publication/wdr2021].

⁶ **Ibid.** p.25-27.

⁷ **United Nations. 2016.** [Data protection regulations and international data flows: Implications for trade and development](https://unctad.org/system/files/official-document/dtlstict2016d1_en.pdf), United Nations Conference on Trade and Development (UNCTAD), Switzerland, p.2, 129. [https://unctad.org/system/files/official-document/dtlstict2016d1_en.pdf].

emphasised the need for data protection. This is because the unprecedented use of data (as was evident during COVID-19) straddles two important areas: the responsibilities of a state towards its citizens and the right of citizens to individual privacy.

In the early days of the pandemic, experts rang out a warning on the “responsible” use of data as “[f]ailing to do so will undermine public trust, which will make people less likely to follow public-health advice or recommendations and more likely to have poorer health outcomes”.⁸ This is one clear and present example of the need for governments to win the confidence of the public in the use of data for overall development. The argument for treating official data as a public good is of recent origin. This comes from the nature of official data and its increasing use by different agents in the society.

This concept of public good is not intrinsic to the data itself. Data are unlike other physical goods whose public provision is seen or sought by all sections of society. Data take different forms, meanings, and uses to different people and cannot be free from the context in which they are utilised. The pandemic and the manner in which governments used personal data to track and control its spread prompted the World Health Organization (WHO) to consider health data as “a public good.”⁹

Although the WHO considers health data as a “public good”, it is also important to factor in the caution expressed by the WDR-2021 that “data are not a pure public good” as they are “excludable”. There are “examples across the public sector of the unwillingness of data holders to share data with other government entities and the public”. Similarly, in the “private sector, firms may not want to sell or exchange their data with others” for reasons relating to either capacity constraints, security issues in sharing data, lack of incentives, or legal constraints.¹⁰

A new social contract

This recognition and utilisation of individual data by governments requires a new understanding between citizens and states. This was reflected in the WDR 2021, when it called for

⁸ Ienca, M., and Vayena, E. 2020. *On the responsible use of digital data to tackle the COVID-19 pandemic*, *Nature Medicine*, March 27, 26, pp. 463–464. [<https://doi.org/10.1038/s41591-020-0832-5>].

⁹ World Health Organization. 2021. *Health Data as a global public good – a call for Health Data Governance 30 September*. September 29. [<https://www.who.int/news-room/articles-detail/health-data-as-a-global-public-good-a-call-for-health-data-governance-30-september>].

¹⁰ In Economics, a Public Good must meet three criteria: non-rivalry, non-exhaustibility, and non-excludability.

“a new social contract for data—one that enables the use and reuse of data to create economic and social value, promotes equitable opportunities to benefit from data, and fosters citizens’ trust that they will not be harmed by misuse of the data they provide. A well-designed data governance framework allows countries to capture the full economic and social value of both public intent and private intent data and leverage synergies between them. This involves creating trust in the integrity of the data systems, while ensuring that the benefits of data are equitably shared.”¹¹

In India, the push towards mainstreaming data for development was expressed in the *Economic Survey* presented in Parliament in 2019, which devoted a chapter titled ‘Data “Of the People, Data By the People and Data For the People”’.¹² Large quantities of data are furnished “by the people” as part of enrolment and participation in beneficiary schemes or while availing public services like health, education etc. These are, to use terminology from the *Economic Survey*, 2018-19, “data of the people” collected by the government, most often with a specific intent known to the respondents. In practice, however, most of these data gathering processes lack explicit voluntary consent of the respondents.

The data ‘by the people’ originate from the day-to-day transactions or interactions of people with other agencies as is the case of ‘digital footprints’ left behind after such transactions. The digital nature of these transactions using standard identifiers such as PAN, Aadhaar, mobile number, email and social media identity makes the collection, scrutiny and processing of such data easier than other data-gathering mechanisms. However, all these forms invariably come with the “lack of explicit consent”.

It would appear that the old paradigms such as the use of data as inputs in evidence-based policymaking or data for development are getting replaced by the new one: ‘data is development’. As India moves to a digital economy, Statisticians are expected to generate enough data that will find huge private markets which require these inputs for effective decision making. For this to materialise, there should be proper data infrastructure in place to capture every need and deed of the citizens. With the cost of building this infrastructure getting cheaper, the talk of data as ‘the new oil’ hinges on the possibility of monetising data by converting it into ‘subscribed’ services.

The above framework is a far cry from the days when country had practically little knowledge of its population, economy and society. The next Chapter looks at how the data collection system evolved over the years.

¹¹ International Bank for Reconstruction and Development/World Bank. 2021. Op. cit. p. xi.

¹² Government of India, 2019. [Economic Survey 2018-19. Volume 1, Chapter 4](https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/vol1chapter/echap04_vol1.pdf), Ministry of Finance, Department of Economic Affairs. New Delhi. Pp. 78-97. [https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/vol1chapter/echap04_vol1.pdf].

III. EVOLUTION OF OFFICIAL STATISTICS IN INDIA

The common people of India are, it may be repeated, anything but statisticians by nature. About such a vital matter as the rent he pays for his land, the ordinary villager is often as vague as he is on the subject of his age. Ask a young peasant how old he is, and he will, as likely as not, reply “Twenty or thirty”; ask his mother her age, and the old lady will hazard “Forty or fifty”; while the grandfather will invert the order and probably tell you that he is “a hundred or eighty.” If you are riding down the countryside and try to get a passing rustic to tell you how far you are from your next camp, he will give you “Eight or ten miles,” and the chances are that you will find it fifteen.”

Lord Meston

President of the Royal Statistical Society
Inaugural Address to the Royal Statistical Society,
November 15, 1932¹³

Lord Meston’s remarks made 90 years ago, in some measure, signal the approach towards data by the commoner, with significant exceptions, even in present day India. This points to an important failing, perhaps not unique to India: that of a prevalence of ‘statistical illiteracy’. This will be dealt with briefly in the final chapter, ‘The Way Forward’.

The history of India’s official statistics before independence is closely linked to the colonial administration’s efforts to record the state of society and economy of the colony, against the backdrop of the near absence of any such systematic records for the country. Most of these came out in the form of Gazetteers at the district or national levels. This is not to ignore the ancient prescriptions for governance in ‘Arthashastra’ or the land administration system initiated during the Mughal period. Most of colonial administrative efforts are meticulously reported in S. Subramaniam’s paper published in the Journal ‘*Sankhya*’ in 1960¹⁴. Subramaniam was with the Directorate General of Commercial Intelligence and Statistics (DGCI&S), the first agency set up for official trade statistics in the country in 1869.

The other key pre-independence developments included the Conference on Agricultural Statistics in 1884, the creation of a Statistical Bureau in Calcutta [now Kolkata] in 1895 and its absorption by a Directorate General of Commercial Intelligence and Statistics (DGCI&S) in 1905, the appointment of

¹³ Meston, L. 1933. [Statistics in India](https://www.jstor.org/stable/2341867), *Journal of the Royal Statistical Society*, 1933, Vol. 96, No. 1, pp. 1-14. [https://www.jstor.org/stable/2341867].

¹⁴ Subramaniam, S. 1960. *A Brief History of the Organization of Official Statistics in India during the British Period*, *Sankhyā: The Indian Journal of Statistics* (1933-1960), Indian Statistical Institute, January, Vol. 22, No. 1/2 (Jan., 1960), pp. 85-118, Indian Statistical Institute.

a Director of Statistics-in 1914, Reports of the Economic Enquiry Committee for 1925, the Royal Commissions on Agriculture and Labour of 1928 and 1930, the Bowley-Robertson Report of 1934, and the Inter-Department Committee on Official Statistics of 1945. The population censuses starting from 1861 need special mention in view of the meticulous care and importance attached to them. All these were important steps intended to contribute to the understanding of the political, social and economic aspects of the country, spanning a period of close to 100 years.

It is interesting to note that land record system whose origin goes back to the Mughal period survives to this day in most parts of the country. There are also cases where the government's control of production and trade of key items such as salt and tea ensured that the concerned administrative authorities maintained records of production, exports and consumption. The Second World War also brought in certain systems for collecting data possibly to aid war efforts.

The developments after the independence of India are briefly noted in the Report of the National Statistical Commission headed by C. Rangarajan¹⁵ that reviewed the statistical system and the entire gamut of official statistics in the country. Free India's government, its commitment to improve the lives of people, and the introduction of Five Year Plans, ushered in a development centric focus to data collection efforts.

The National Income Committee, appointed in 1949, and published its report in 1951¹⁶, highlighted key gaps in the statistical database of the country, but "decided to discuss these problems in greater detail before making definite recommendations".¹⁷ In the following years of independent India, exclusive statistical offices like the Central Statistical Unit which later converted into the Central Statistical Organisation (CSO, 1951), the Directorate of the National Sample Survey (1950) and the Computer Centre (1967) were established.

The National Income unit that functioned under the Ministry of Finance was transferred to the CSO in 1954. The Indian Statistical Institute (ISI) set up by the late P.C. Mahalanobis in 1932 at Kolkata grew into an institution of international repute. Path-breaking research in statistical theory and practice by the

¹⁵ **Ministry of Statistics and Programme Implementation. 2001.** [Report of the National Statistical Commission](https://www.mospi.gov.in/web/mospi/report-of-dr.-rangarajan-commission), Government of India, September 5. [https://www.mospi.gov.in/web/mospi/report-of-dr.-rangarajan-commission].

¹⁶ **Department of Economic Affairs. 1951.** [First Report of the National Income Committee, April 1951](http://14.139.60.153/bitstream/123456789/9267/1/First%20Report%20of%20The%20National%20Income%20Committee%20April%201951_HK_PCL8737.pdf), Ministry of Finance, Government of India, April. [http://14.139.60.153/bitstream/123456789/9267/1/First%20Report%20of%20The%20National%20Income%20Committee%20April%201951_HK_PCL8737.pdf].

¹⁷ **Ibid.** p. 51.

ISI provided direct inputs to the development of the official statistical system. This came in the form of pioneering work on large-scale sample surveys, design of agricultural experiments, statistical quality control, economic planning, and use of electronic computers in statistical work.

Renowned experts in Statistics and Economics were associated with these developments and the efforts to improve the statistical system gave the country a head start among developing countries. It also placed India in the forefront of leadership in statistical theory and practice internationally.

Official statistics in a federal framework

The Indian federal structure influenced the organisation of the statistical system as well. The division of administrative functions between the Government of India and the State Governments on the basis of the Union, State, and Concurrent Lists also determines the roles and responsibilities of statistical organisations. At the centre, the responsibilities are further divided amongst the various ministries and departments, according to the Allocation of Business Rules. The collection of statistics on any subject generally vests in the authority (Union Ministry or Department or State Government Department) responsible for that subject according to its classification under Seventh Schedule of the Constitution. This system assumes that the flow of statistical information originates from the States to the Centre except in cases where the State-level operations are an integral part of centrally sponsored schemes or data are collected through national sample surveys.

The collection of statistics for different subject-specific areas, like agriculture, labour, commerce, industry, etc. vests with the corresponding administrative ministries. Some of the Union ministries, for instance, Agriculture, Water Resources, and Health have full-fledged statistical divisions, while most others have only a nucleus cell. Large-scale statistical operations like the Population Census, Annual Survey of Industries, Economic Census, are generally centralised. More about this later.

The statistical systems in the States are similar in structure to that at the centre. It is generally decentralised laterally over the Departments of the State Government, with major Departments, such as Agriculture or Health, having large statistical divisions for their work. At the apex level is the Directorate of Economics and Statistics (DES), which is formally responsible for the coordination of statistical activities in the State. The DESs have organisations at the State headquarters, with statistical offices in the districts and taluks. The statistical activities of the DES in various States are more or less uniform. They publish statistical abstracts and handbooks, annual economic reviews or surveys, district statistical

abstracts, and state budget analysis, estimates of the State Domestic Product and Retail Price Index Numbers, and engage in other statistical activities specific to the State. A cursory look at their outputs shows that even in the current digital age most of these are available only in the form of print publications with very few in open digital formats.

IV. GROWING CONCERN OVER DATA INADEQUACIES IN INDEPENDENT INDIA

“The operational efficiency of the Indian Statistical System today is compromised by serious deficiencies with respect to credibility, timeliness and adequacy.”

Report of the National Statistical Commission

(Paragraph 1.1.2)

While the pre-independent initiatives are more of historical interest, the developments after independence need to be looked into in more detail. As these institutions and processes are still in place and the manner in which these are coping with the new data architecture is critical to understand the present state of the official statistical machinery.

From another perspective – that of the role played by statistics in state policies – there is a difference between why data are collected by a colonial regime and a sovereign democratic republic:

“[T]hroughout the British period the statistical development was geared towards administration, trade, commerce and such other activities. It is only after the independence in 1947 that the country saw an urgent need for a statistical framework suitable for economic and social development”¹⁸.

One of the key features of the newly created institutions and systems in independent India was the nodal role played by the central agencies in their formulation and implementation. Given that the States were first formed through integration of the British Presidencies, Princely States and French and Portuguese colonies with the Union, and their subsequent re-organisation along linguistic lines, it was clearly not expected of them to develop these systems on their own.

The beginning of the planning era also required data at the Union and State levels without time delay. Thus, the central statistical bodies and the Ministries were the ones that initiated important surveys and censuses. In areas such as agriculture, where the States had the key role, schemes to improve data reporting were centrally sponsored. However, most of the systems were built along the style of other administrative initiatives with conventional style bureaucracy where the entire statistical machinery was subordinate to the administrative hierarchy.

¹⁸ Ghosh, J.K., et al. 1999. *Evolution of Statistics in India*, *International Statistical Review/Revue Internationale de Statistique*, Vol. 67, No. 1, April, pp.13-34. International Statistics Institute. [<https://www.jstor.org/stable/1403563>].

The exceptions were the National Sample Survey (NSS) and the Central Statistical Organisation (CSO), created in 1950 and 1951, respectively, which enjoyed a great degree of autonomy. While the former was to serve as “a multi-faceted fact-finding body”, the latter’s objective was “to coordinate the statistical activities of independent India”, and among others, to “prepare and publish” monthly and annual Statistical Abstracts, “to act as a liaison with [the] United Nations Statistical Office and to disseminate annual statistics by graphs and charts as well as tables for public use”¹⁹.

The Computer Centre was started in 1967 “as an office attached to the then Department of Statistics, Cabinet Secretariat, to cater to the data processing needs of the Statistics Department and other Departments in the Government of India”²⁰. It was one of the first computer centres anywhere in the Government and possessed the latest machines then available.

Initially, India’s statistical system was not subject to critical scrutiny, though it was always felt that it was not keeping pace with the data needs, especially those required by the Planning Commission. The Planning Commission had all along given a helping hand to the statistical system as a responsible care giver. Mahalanobis, who guided the statistical system also played a key role in the formulation of the Five Year Plans in the early years of the planned economy. Nikhil Menon (2022) brings out the role played by the Mahalanobis and the ISI in the evolution of the economic planning²¹.

Futility of chasing administrative data

The administration of various statutes were expected to generate large quantities of data as a by-product. Most statutes had provisions for submitting periodical reports by the agencies using the powers vested in these laws. However, by the 1980s, it was clear that the administration of most of these statutes could not produce timely and complete reports. For example, the setting up of employment exchanges all over the country and the prescription that all employment and vacancies were to be reported by organisations and all unemployed looking for jobs expected to register themselves in these exchanges was expected to produce the necessary statistics for employment and unemployment²². In reality, however, the statistics

¹⁹ **Ibid.**

²⁰ **Ministry of Statistics and Programme Implementation. 2001.** *Report of the National Statistical Commission.* Government of India. p. 447.

²¹ **Menon, N. 2022.** *Planning Democracy: How a Professor, an Institute, and an Idea Shaped India.* Penguin Viking. India.

²² The Employment Exchanges (Compulsory Notification of Vacancies) Act was enacted in 1959 to provide for compulsory notification of vacancies to the Employment Exchanges and for the rendition of returns relating to Employment situation by the employers. This act came into force with effect from May 1, 1960.

coming out from this process were highly incomplete and failed to measure the actual employment situation in the country. There were also inevitable delays in bringing out official statistics considering the size and complexity of the country – an example of a critical gap in official data that has direct impact on policy making. There are many other examples of the failure of administrative systems as a source of official statistics that persist.

Pandemic exposes the chronic malaise

The COVID-19 pandemic brought to the fore the limitations of the Registration of Births and Deaths Act, 1969, and its amendments to provide timely statistics on the pandemic-related deaths. The failure of the official machinery to reflect accurately the enormity of human loss as a result of the COVID-19 pandemic was not an entirely new revelation. Two decades earlier, the Rangarajan Commission had sounded a note of warning in its Report:

‘over the years, the Administrative Statistical System has been deteriorating and has now almost collapsed in certain sectors. The deterioration had taken place at its very roots namely, at the very first stage of collection and recording of data, and has been reported so far in four sectors: agriculture, labour, industry and commerce. The foundation on which the entire edifice of Administrative Statistical System was built appears to be crumbling, pulling down the whole system and paralysing a large part of the Indian Statistical System. This indisputably is the major problem facing the Indian Statistical System today²³.

Although the civil registration system was introduced long back during the colonial period, registration was voluntary, different provinces had different legislations, and there was no standardisation of concepts, definitions and classifications. The enactment of the ‘Registration of Births and Deaths (RBD) Act, 1969’ replaced the diverse laws that existed and brought a uniform legislation in the country. The Act provided for a hierarchical set-up for the registration starting at the panchayat level.

As noted by the Rangarajan Commission, a combination of administrative factors is responsible for the poor registration levels of vital events. Except for a few States and UTs, generally multiple agencies are involved in the registration work at the sub-national level. Moreover, for the functionaries at all levels, the work related to registration of births and deaths is in addition to their other normal duties, and is generally performed in an honorary capacity. In addition to possible administrative apathy, other factors that lead to incomplete registrations are a general lack of public awareness about the statutory

²³ Ministry of Statistics and Programme Implementation. 2001. Op. cit. p.450.

requirements and procedures of registration, lack of demand of birth and death certificates in rural areas, and the perception that there is nothing to gain from registration. Computerisation of the registration work has substantially reduced the reporting the delays. The requirement of providing birth certificate for school admissions and death certificates for settling claims after death of an individual however ensures registration process will eventually reach perfection in the coming days.

That said, even in States with 100 per cent civil registration, its use by local authorities in their work is problematic. For instance, even now almost all local authorities report the population in their jurisdiction based on the 2011 Census because of procedural difficulties in changing the information gathered at the time of census collection, specifically the 'present address of usual residence' and other vital events.

Even earlier, the Registrar General of India had initiated a Sample Registration System (SRS) from the 1960s to provide reliable estimates of birth and death which now includes medical certification of cause of death as well. The SRS was introduced because the birth and death registration system was not able to provide reliable vital statistics. Although based on a sample, the SRS is still considered to be far superior to the universal registration system as it is based on sound statistical methodology with a strong supervisory mechanism for data quality and coverage.

The main objective of the SRS is to provide reliable estimates of birth rate, death rate and infant mortality rate at the national and state level. It consists of a base line survey of the sample units to obtain demographic details of the usual resident population of the sample areas and a continuous (longitudinal) enumeration of vital events of usual resident population. There are independent retrospective half-yearly surveys for recording births and deaths which occurred during the half year under reference and updating the population details. Further matching of events recorded during continuous enumeration and those listed in course of the half-yearly survey, and a field verification of unmatched and partially matched events ensure complete recording of all vital events.

The large scale nationwide sample surveys or the NSS was a novel arrangement initiated under the aegis of the ISI, Kolkata, for meeting national data needs in economic and efficient way. The NSS was reorganised in 1974 and all activities, including those performed by the ISI, were brought under one umbrella. This was expected to bring speed and efficiency in survey operations and report generation. The NSS was expected to remedy the weakness of administrative statistics in many areas and to this extent its surveys covered a large number of topics. However, despite the reorganisation, it failed to bring out its reports on time and substantial parts of the data that were collected remained unanalysed.

Hesitancy on the part of the statistical agencies, including the CSO and NSS, to absorb or adopt modern computing technologies – without which they would not be able to meet the emerging data needs – is one of the reasons for this lag in data collection and finalisation of reports. In the 1980s, the National Informatics Centre (NIC) took over the responsibility of integrating data from different Ministries and providing computing resources for most of the data processing activities of the Government. The role played by the NIC will be discussed later in this report.

Failure of the coordination mechanism

The CSO has a major responsibility to coordinate statistical activities across domains. This, as noted by the Rangarajan Commission, depended on two factors: the degree of its initiative and ability to persuade various ministries and departments to share data, and the co-operation of the ministries, in particular, their willingness to participate in this process as a team and to be persuaded to accept the conclusions of the team about their statistical work.

However, given the historical background, India's statisticians had a less flexible mode of thinking, averse to change and 'outside influence'. An example of this lies in an earlier attempt to infuse responsibility and oversight in the statistical mechanism. As early as the 1980s, the unsatisfactory experience in Indian statistical system in coordinating with the Ministries led the CSO to search for an institution outside of itself and the ministries from which it could derive authority.

The idea was concretised by the Committee to Review the National Statistical System (1980), also referred to as the Kripa Narain Committee, in two of its recommendations²⁴. One required the Government of India to formally declare by Executive Order that the Department of Statistics, to which the CSO belonged, was the "Nodal" department for undertaking integration of data required for Government's decision making, for setting and maintaining standards, and for improvement and development of statistics in all respects. The other was to create a National Advisory Board on Statistics (NABS) with the Deputy Chairman or Member-in-charge for Statistics of the Planning Commission as the Chairman and the Director General of CSO as the Vice-Chairman. The Government constituted NABS in 1982. From 1992, the NABS was chaired by the Minister in charge of Statistics. However, this Board turned out to be ineffective, primarily because of lack of official or legal support.

²⁴ **Department of Statistics. 1980.** [Report of the Committee to Review the National Statistical System](http://14.139.60.153/bitstream/123456789/1832/1/Report%20of%20the%20Committee%20to%20Review%20the%20NSS%20June%2c1980%2c%20Department%20of%20Statistics%2c%20Ministry%20of%20Planning%20Govt%20of%20India.D8728.pdf), Ministry of Planning, Government of India, June, p.97. [http://14.139.60.153/bitstream/123456789/1832/1/Report%20of%20the%20Committee%20to%20Review%20the%20NSS%20June%2c1980%2c%20Department%20of%20Statistics%2c%20Ministry%20of%20Planning%20Govt%20of%20India.D8728.pdf].

One available institutional arrangement for coordination was the Conference of Central and State Statistical Organisations which provided a forum for the States and central agencies to interact once a year. This mechanism, vibrant till the 1970s, nearly stopped functioning until it was revived in 2000 after the NSC started functioning.

V. REVIVING INTEREST IN OFFICIAL DATA

“The need for timely and reliable statistics for policy formulation and planning cannot be over emphasised. There is reason to believe that with progressive dismantling of the system of economic controls, the quality of data flows has weakened. Government has decided to establish a National Statistical Commission to critically examine the deficiencies of the present statistical system with a view to recommending measures for a systematic revamping of the system”.

Yashwant Sinha

(former Union Finance Minister)

Union Budget speech, 1999²⁵

The genesis of the Rangarajan Committee can be traced to this Union Budget Speech on February 27, 1999, by the then Finance Minister, Yashwant Sinha. These reforms in the official statistics sector were proposed as the economic reforms initiated in the early 1990s were expected to widen the scope of data for the user community.

The setting for this announcement to establish a National Statistical Commission (NSC) was also due to the increasing integration of Indian economy with the global economy, putting pressure on government to take a hard look at the official data systems. An additional need came from international agencies to improve the compilation of economic data by member countries. Multilateral bodies like the IMF and World Bank started taking a keen interest in developing statistics and their timely reporting by member countries especially after the East Asian crisis of Nineties through the introduction of the Special Data Dissemination Standards (GDSS) and General Data Dissemination Standards (GDDS) for statistical reporting.

The key recommendations by the Rangarajan Commission Report submitted in 2001, included the setting up of a “permanent and statutory National Commission on Statistics”²⁶ and the reorganisation of a State’s statistical activities under their respective Directorates of Economics and Statistics, with the latter to be headed by “a professional statistician or a professional economist”.²⁷

²⁵ Sinha, Y. 1999. *Budget 1999-2000*. (Speech made at the Lok Sabha.) Ministry of Finance, Government of India, February 27. [<https://www.indiabudget.gov.in/budget2021-22/doc/bspeech/bs19992000.pdf>].

²⁶ Ministry of Statistics and Programme Implementation. 2001. Op. cit. p. 456.

²⁷ Ibid. p. 469.

Acting on this report, in June 2005, the Government set up the NSC with a part-time Chairman and Members and an entity called National Statistical Organization (NSO) with the National Sample Survey Organisation (NSSO) and the CSO as two separate wings of the NSO. The NSO was to act as per the policies and priorities set by NSC. The Chief Statistician of India, in the rank of Secretary to the Union government, was to head the NSO and also to function as Secretary of the Commission. Oversight of the technical functioning of the NSSO was also transferred to the NSC from the NSSO's Governing Council, which was dissolved.

These measures gave hope that the Indian statistical system would transform into a modern entity enjoying autonomy and national importance. There were similar efforts in countries like UK to reform their statistical system through structural legislation that would insulate official statistics from political considerations and bring in technical considerations alone in the functioning of the statistical system.

The NSC was first set up through a cabinet resolution with the idea that it would become a statutory commission in due course. Despite the Suresh Tendulkar Committee's recommendations in 2008, this was not acted upon. Subsequently, in 2011, a committee under the late Madhava Menon submitted another blueprint to the Commission, including a Draft Bill²⁸. This exercise too did not succeed. Thus, no visible efforts have been made till now to empower the Commission by providing a statutory framework for its functioning.

The Government in an order dated May 23, 2019, restructured the Statistics wing of the Union Ministry of Statistics and Programme Implementation (MoSPI) with the objective to streamline and strengthen the nodal functions of MoSPI and to integrate the administrative functions²⁹. For one, the order created doubts among experts as to how the NSC will play its role and set policies and priorities for the NSO when all its divisions are integral part of the Ministry. Second, the concerns were more in the case of the NSSO which, since its inception, had external technical oversight and catered to the data needs of not just the CSO but of other Ministries as well, and took care of the needs of researchers³⁰. The Government was also ambivalent on the functional status quo and how the independence and autonomy of the CSO

²⁸ **National Statistical Commission. 2011.** *Report of the Committee on Legislative Measures in Statistical Matters as Adopted by the National Statistical Commission*, National Statistical Commission Secretariat, October. [http://164.100.161.63/sites/default/files/committee_reports/legislative_measure_stat_matter_18jan12.pdf].

²⁹ **Press Trust of India. 2019.** *MOSPI orders revamp, merges NSSO, CSO into NSO*, The Times of India, May 25. [<https://timesofindia.indiatimes.com/business/india-business/mospi-orders-revamp-merges-nso-cso-into-nso/articleshow/69496878.cms>].

³⁰ **Shetty, S.L. 2020.** *Indian Statistical System in a Troubled State*, Economic and Political Weekly, Vol. LV N 3, January 18. [<https://www.epw.in/journal/2020/3/perspectives/indian-statistical-system-troubled-state.html>].

and the NSSO will be protected under the new organisational arrangement. A subsequent clarification³¹ does not appear to have resolved the matter as vacuum continues to persist at the leadership of the national apex statistical body, and the efforts initiated by the Rangarajan Commission to isolate the statistical system from Government have lost their focus.

The need to insulate the statistical system from Government influence is a fundamental requirement to ensure its autonomy. In addition, the enormous advancements in the field of data collection, processing, and dissemination call for statistical organisations to be more nimble than conventional government bureaucratic agencies.

As is the norm nowadays, when the UN's Sustainable Development Goals have become national and global benchmarks, the need to monitor these outcomes and absorb new IT-based solutions to meet data gaps to ensure that India's data architecture is synchronised with international releases of indicators also call for statistical agencies to be dynamic and innovative in their attitude and approach. Somehow, however, India's statistical mechanisms appear to be caught in a time warp. The 'latest' data made available for most indicators inevitably come with a lag of a few years, if not more. Alternatively, they are projections, which, howsoever corrected using statistical techniques, will not be as reflective as the latest primary data that have been analysed and released on time.

Although the efforts to restructure India's statistical system started in 1999, more than 20 years later it is not very far from where it started. The more things change the more they appear to be the same. However, now there are many models available to ensure that India can join the list of nations that have clear and usable data. As noted earlier, the UK Statistics Authority which describes itself as “an independent body at arm's length from government”³² is a statutory office that oversees the functions of the Office of National Statistics that produces almost all key national statistics. The authority also does independent monitoring and assessment of official statistics through a Code of Practice for Statistics, which “sets standards that producers of official statistics should adhere to”. The Office of Statistics Regulation, an independent regulatory arm of the UK Statistics Authority, “assesses compliance of these statistics against the Code of Practice for Statistics”³³.

³¹ **Ministry of Statistics and Programme Implementation. 2019.** *Restructuring of Ministry of Statistics & Programme Implementation (MoSPI)*, Press Information Bureau, Government of India. [<https://archive.pib.gov.in/newsite/PrintRelease.aspx?relid=190141>].

³² **UK Statistics Authority. 2022.** *What we do*. [<https://uksa.statisticsauthority.gov.uk/what-we-do/>].

³³ **UK Statistics Authority. 2022.** *Code of Practice for Statistics – Ensuring official statistics serve the public*, (Edition 2.1, Code revised on May 5, 2022) [<https://code.statisticsauthority.gov.uk/wp-content/uploads/2022/05/Code-of-Practice-for-Statistics-REVISED.pdf>].

VI. DATA AS A POLITICAL TOOL

“In the original sense of the word, ‘Statistics’ was the science of Statecraft; to the political arithmetician of the eighteenth century, its function was to be the eyes and ears of the central government. It could tell the Prince how many able-bodied men he might mobilise, how many would be needed for the essentials of civil life; how numerous or how wealthy, were sectarian minorities who might resent some contemplated change in the laws of property, or of marriage; what was the taxable capacity of a province, his own, or of his neighbours.”

R.A. Fischer

Presidential Address to the First Indian Statistical Congress, 1938³⁴

Data has found three major uses in India’s political environment. In governance, they are key indicators to set targets and monitor progress. During electioneering, contestants cherry-pick information either to woo voters or to run down their electoral opponents. For researchers, journalists and others contributing to the enhancement public knowledge, data – both public and private – come in handy to either support their arguments or to counter opposing narratives.

These three political uses of data are not exclusive of each other and at times, with official data become a matter of public debates. One such debate, in which official data entered the public consciousness through a series of front page newspaper reports, and across visual and online media was the controversy over two leaked reports NSSO – one of which was released later and the other withheld due to ‘quality issues’.

The first of these two NSS reports was on the employment-unemployment situation in the country based on the first Periodic Labour Force Survey (PLFS) slated for release in December 2018. This was the first national survey on employment-unemployment by NSSO after 2011-12. As per the established procedure, NSS reports are released after approval by the NSC. However, although this report was approved by NSC for release in December, it was not released. A leaked version of the report was published by a journalist in the daily, Business Standard, after which the NITI Aayog came out in defence of the withholding of the report. The same report was, however, released by the NSS a few months later by which time Government had got elected afresh. A few months later, another NSS report on household consumer expenditure, a report that forms the basis for estimating poverty numbers, was

³⁴ Fisher, R. A. 1938. *Presidential Address*, *Sankhyā: The Indian Journal of Statistics* (1933-1960), 4(1), pp. 14–17. [<http://www.jstor.org/stable/40383882>].

published by the same journalist. Once this report was leaked, the Government came out with a statement mentioning ‘quality issues’ for abandoning the survey altogether without any comment of on what these quality issues are. The leaked version of the report showed a much bleaker picture of poverty as there was hardly any improvement in the household expenditure since the previous survey in 2011-12.

Before this, there were debates over the new GDP series and the back-series computation. Changing the base year for GDP computations is a standard practice to accommodate methodological innovations, availability of new data sets, uniformity with other indicators etc. The CSO changed the base year from 2004-05 to 2011-12 and the new base year was adopted from 2015-16. For time series comparisons, the usual practice is to prepare a GDP back-series using the methodology and data sets as used in the new base year. Back series preparation was problematic for the new series due to non-availability of corporate data for the earlier years. The NSC had got an exercise undertaken by experts to prepare the back-series. However the CSO prepared a back series that produced a different picture of the economy compared with the NSC exercise. This could be explained in view of the different methodology used in the two exercises. The controversy arose due to the fact that the new back series was released by the NITI Aayog instead of the CSO. This clearly brought out the political interest indirectly influencing data based narratives. It is important, therefore, to look at the relationship between statistical agencies and the government in the light of these debates.

Official statistics and the Government

Some of these controversies drew several commentaries from economists, the media, and the political establishment on the integrity and autonomy of Indian official statistics. The official reactions came from NITI Aayog, a generalist body created as a diluted version of the Planning Commission, rather than the specialist statistical agencies involved – a visible subordination of these agencies by the political establishment who have a stake in creating a particular narrative.

The argumentative nature of the Indian polity, the widespread access to independent media and the failure of the Government to address what in effect are technical issues made the debate a novelty for the country’s public. This was not the case at the international level, where there are examples of governments dealing with uncomfortable statistics in a variety of ways. A publication by the Economic Governance Support Unit of European Parliament, *Statistical Governance in Greece – Recent Developments*, in November 2016, deals with certain developments in the Greek statistical system. It also covers legal proceedings before Greek courts against the Head of the Greek Statistical Office, ELSTAT during 2010

to 2015, among others on the ground of the accusation that he inflated the 2009 budget deficit for “undermining the national interest.”³⁵

It is well known that one of the major causes of the collapse of the Greek economy was attributed at that time to the frequent revisions of official statistics. The frequent revisions made them unreliable according to the inspection reports of the Eurostat, the European agency overseeing the official statistics for European Union. Although prosecuting the chief statistician for following the established procedure of compiling official data on the ground of falsifying numbers, as these were not to the taste of the government, was a strange and extreme case, it evocatively brings out the centrality of official data.

William Seltzer in a report prepared for the UN Statistical office (UN, 1994) deals with many possibilities of politics coming in to play in official statistics. The report quotes the Norwegian Central Bureau of Statistics (CBS) as saying:

“...CBS tries to avoid provocative comments relating to issues which are hotly debated between our political parties. However the CBS will (as a policy) not refrain from putting its statistics in a relevant social or economic context, irrespective of how this will be interpreted by various parties.....**The CBS will show circumspection, while Ministries and Ministers occasionally will have to live with some statements from the CBS that they do not like; extreme caution would lead to a reduction in the social usefulness of the CBS** [emphasis by author].”

This is another side.

Coming to the Indian context, it is instructive to see how data entered the jobs narrative. In an interview to the magazine, Swarajya, the Prime Minister was quoted as saying “...On this issue, more than a lack of jobs, the issue is a lack of data on jobs. Our opponents will naturally exploit this opportunity to paint a picture of their choice and blame us”. He went on to give more details to indicate where and how jobs are being created:

“...If we look at numbers for employment, more than 41 lakh formal jobs were created from September 2017 to April 2018 based on EPFO payroll data. According to a study based on EPFO data, more than 70 lakh jobs were created in the formal sector last year..... In just one year, 48 lakh new enterprises got registered. Will this not result in more formalisation and better jobs? More than 12 crore loans have been given under

³⁵. **European Parliament. 2017.** *Statistical governance in Greece – recent developments* [Briefing], November. [[https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/614481/IPOL_BRI\(2017\)614481_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/614481/IPOL_BRI(2017)614481_EN.pdf)].

Mudra (micro loans). Is it unfair to expect that one loan would have created or supported means of livelihood for at least one person? More than one crore houses have been constructed in the last one year; how much employment would this have generated? If road construction has more than doubled per month, if there is tremendous growth in railways, highways, airlines, etc., what does it indicate?”³⁶

Quoting sources with poor statistical legitimacy as evidence of rise in employment could be due to the absence of independent data from labour force surveys or their weakness in countering the claims. However, these ‘leading’ and ‘loaded’ questions coming from the Prime Minister can put the national statistical agencies involved in measuring employment under extreme caution.

In the past, debates on data – whether on GDP, poverty, or employment – were mostly confined to the technicalities of the data generating processes and the interpretation. This was understandable given that access to reports and data was limited and the tools of analysis and dissemination poor. The heightened interest of print and digital media in official numbers is somewhat a new phenomenon. Added to this is the instant circulation of information through social media. The vast pool of bright researchers with data analytics tools at their fingertips now provides abundant scope for discussions and instant publication of findings. These challenging scenario needs to be fully integrated in to the statistical system. The official agencies however are not trained to tackle this new scenario. The most unfortunate fallout of these changes has been that the agencies withdrew into a shell; adopting a wall of caution in collecting and releasing data.

³⁶ Jagannathan, R. 2018. *Swarajya Interviews Prime Minister Modi – Part I: The State of Indian Economy*, Swarajya, July 3. [<https://swarajyamag.com/economy/swarajya-interviews-prime-minister-modi-the-state-of-indian-economy>].

VII. THE FEDERAL DISCONNECT

“The States under our Constitution are in no way dependent upon the Centre for their legislative or executive authority. The Centre and the States are co-equal in this matter.”

B.R. Ambedkar, November 25, 1949

Concluding remarks in the Constituent Assembly on the Constitution

A key feature of governance in India is the sharing of powers and responsibilities under the Seventh Schedule, which provides for Union, State and Concurrent Lists. The Chairman of the Constitution’s Drafting Committee, B.R. Ambedkar, in his concluding remarks in the Constituent Assembly, highlighted the significance of these three lists and the sharing of powers.

“The chief mark of federalism, as I said, lies in the partition of the legislative and executive authority between the Centre and the Units by the Constitution. This is the principle embodied in our constitution. There can be no mistake about it. It is, therefore, wrong to say that the States have been placed under the Centre. Centre cannot by its own will alter the boundary of that partition. Nor can the Judiciary.”³⁷

Ambedkar’s formulation of separation of powers faces a challenge when it comes to official statistics because of overlapping domains and administrative interdependence between the Union and the States. It was mentioned earlier that most key statistical initiatives originates from the centre, even where the subject matter actually fell in to the State List. Under this sharing of powers, ‘Census’ and ‘Inquiries, surveys and statistics for the purpose of any of the matters in the Union List’ come under the exclusive domain of the centre. The Concurrent List includes ‘vital statistics including registration of births and deaths’ and ‘inquiries and statistics for the purposes of any of the matters not specified in state list and concurrent list. Given that definition of what constitutes the statistics on any matter is somewhat blurred, in reality there appears to be no hard and fast rule on the agencies that can collect data on any specific data in the country and publish it.

Given the structure of India’s polity, where States are the implementing agencies, even for centrally funded schemes, a clear understanding between the Union and the States, backed by a mechanism for data collection, analysis, publication, and dissemination, gains critical importance. The “Census”, which

³⁷ **Ambedkar, B.R. 1949.** Dr. B.R. Ambedkar’s Concluding remarks in the Constituent Assembly on Constitution on November 25, 1949. In *B.R. Ambedkar, Selected Speeches*. Prasar Bharti. pp. 39-40.

is listed under the Union List gives a clear example of the need for the Union and the States to work together for the creation of a credible, reliable, and timely statistical data architecture.

Most of the data are required at the national level with disaggregation at the State or possibly district levels. Therefore, it is seen as essential for the Centre to carry out all pan-India statistical operations that include all Censuses. This creates a structure ensuring adoption of uniform definitions enabling comparability and aggregation. It is possibly this aspect that has made the States take a back seat in statistical initiatives, as they would have to follow what the Union ministries set out as guidelines.

This current administrative mechanism raises several issues. In many sectors like agriculture, education, and health to name a few, one would expect the data to be generated at the field or at the local institutional level and aggregated to successive higher domains and produce national level estimates. But inter-State variations in statistical capacity and administrative efficiency create deficiencies in the basic building block of the statistical framework – data collection. We thus have the respective central ministries playing a major role in guiding and directly involving in the collection and dissemination of statistics on these topics. One cannot, therefore, fault the Central agencies in exercising control of data generation and dissemination. The flip side of this arrangement has been the decline of the State's role in the official statistical system.

Estimating the Gross State Domestic Product (GSDP) for the State is one of the important activities for DES, which is a State body. The GSDP has assumed critical importance in recent times as the successive Finance Commissions have tied the devolution of funds and the borrowing limits to the GSDP. As with national GDP, estimating the GSDP is a complex exercise and calls for data from all economic agents including households, government, corporates, and non-profit organisations.

A standard national prescription is not always possible given the diverse economic structure of States. For a large number of sectors controlled by the Central government, such as the railways, air transport, and post and telecommunication, the States have to depend on the CSO. In the latest revision of the GDP, with 2011-12 as the base year, the CSO started using the corporate financial data from the corporate filings with the Ministry of Corporate Affairs. The new GDP series, for the most part, replaced the Annual Survey of Industries with corporate financial data as the source for estimating manufacturing value added. This data is not available at the State level and this procedure has meant that the States have to look to the CSO for data for their GSDP compilation. This creates not only a complex process, but further restricts the States – where all the economic activities take place – from gaining clear insights on

their economic activities, adding to the constraints on States to chart economic policies that are specific to their needs.

State participation in NSS

A key feature of the NSS since its inception was the replication of its surveys on separate and independently selected samples in all States. The idea was that the State's sample will augment the central sample and help the pooling of data to generate sub-State level estimates. This did not materialise as barring a few, States failed to process their data. However, the Governing Council that managed the NSS had representation from the States ensuring the selection of survey subjects also kept in mind the interest of the States. The recent approach appears to be that the issues that are to be surveyed are selected by NSS without any consultations with States and their participation is not desired in most surveys.

Digitisation of reporting

Over the years, several schemes for the benefit of individuals or households have been initiated as 'mission mode projects', which generate information at the household/individual/local government levels. Infusion of IT in monitoring allows the data to be directly fed in to a central server that can generate instant web enabled data points available as Dash Boards. Clearly the custodianship of all these data rests with the central agency in charge of the project. These dash board statistics often come with limited meta-data and the unavailability of the underlying database for research purposes limits their utility.

VIII. INSTITUTIONAL INADEQUACIES: CAUSES AND IMPLICATIONS

“An institutional gap that has been highlighted recently is the need for a strong national statistical system, independent of control by the government. India has a high reputation in this area, but this has been eroded...If we want the economy to appear attractive to private capital – domestic and foreign – then the Indian statistical system must meet the highest standards observed in emerging market countries.”

Montek Singh Ahluwalia

*Backstage – The Story Behind India’s High Growth Years, 2020*³⁸

Statistical frameworks for any country require a strong institutional foundation. In addition to the overlap between the Union and the States in the administrative process, a new impediment to professionalism in the official statistical machinery is overt interference in the release of official data, exemplified by the pole position of the NITI Aayog (the successor body to the Planning Commission, albeit with a narrower remit) in the recent data narratives.

For instance, in 2017, the revised GDP estimates were released in the presence of NITI Aayog officials and it was the same officials from NITI Aayog who defended the Government when the employment report was leaked. The entry of NITI Aayog is a pointer to the direct control over data by a body that is a creation of the government of the day.

The reforms suggested by the Rangarajan Commission wanted a professional as the head of the statistical system and the NSC to function as an independent body to play a nodal role which ought to have been brought in to the picture. These aims, as discussed earlier, have not been met and continue to remain elusive.

The NITI Aayog now has a ‘flagship’ initiative to develop a National Data and Analytics Platform (NDAP) within it to improve access and use of government data. It is expected to be a user-friendly web platform that aggregates and hosts datasets from across India’s statistical system. According to the objectives it has set, NDAP seeks to ‘democratise data delivery by making government datasets readily accessible, implementing rigorous data sharing standards, enabling interoperability across the Indian data

³⁸ Ahluwalia, M.S. 2020. *Backstage: The Story behind India’s High Growth Years*, Rupa Publications India, New Delhi. p. 408.

landscape, and providing a seamless user interface and user-friendly tools'. This is indeed a laudable initiative and the programme should become a major data initiative with the resources and access to high-end IT infrastructure that the organisation has at its command.

One may remember that the Government of India, through the Department of Science and Technology had released a National Data Sharing and Access Policy (NDSAP)³⁹ with somewhat similar ambitions. The preamble of this Policy issued in 2012 stated:

'data collected or developed through public investments, when made publicly available and maintained over time, their potential value could be more fully realized. There has been an increasing demand by the community, that such data collected with the deployment of public funds should be made more readily available to all, for enabling rational debate, better decision making and use in meeting civil society needs'.

The NDSAP covered all types of data including geo-spatial data over which the government had absolute control till then. Under this Policy, each organisation was to share all shareable datasets in an open data format. The technology backup for this is provided by the NIC under Department of IT. The open data portal data.gov.in has been successful in bringing together close to a half a million data resources under different catalogues. It is a state-of-the-art data portal, highly interactive and allowing several API for public use. One might say that NDSAP is India's response to the global Open Government Data movement.

Not to be left out, the MoSPI also developed a National Integrated Information Platform (NIIP) which has uploaded several survey data of NSS providing online retrieval facility of survey estimates by domains and offering limited cross tabulations, including graphic visualisations. These efforts to democratise data through technological innovations are extremely important given that the statistical agencies were rather slow in assimilating IT. As noted earlier, most programme-implementing Ministries now have web-based data capturing mechanisms and dynamic dashboards. The past experience, however, has been that many such initiatives are supply driven and are not flexible enough meet changing user needs and lean towards showcasing government achievements rather than enabling a critical understanding of the reality.

However, the disproportionate efforts to showcase data on the back of modern IT architecture without strengthening the underlying data generation processes raise questions. The Economic Census is one

³⁹ National Data Sharing and Accessibility Policy-2012 (NDSAP-2012), Department of Science & Technology, Ministry of science & Technology, Government of India

such case in point. This is mammoth exercise to list all non-agricultural establishments in the country with the ultimate objective of creating a directory of business activities which can be used as a register for surveys. Until now these were conducted with the help of State Statistical offices. It did help in understanding the growth and spread of economic activities but failed in the basic objective of creating a business register in the face of highly defective data gathered by the enumerators, mostly poorly trained unemployed graduates, anganwadi workers, etc. In spite of the past experience, the Government went ahead with the latest survey mobilising the Common Service Centres (CSC) spread all over the country and the high end technology that the CSC has at its command.

Though a lot of time has lapsed since the Seventh Economic Census which was launched in 2019, the results have not been released. The MoSPI explained away the reason for the delay on State/UT governments not providing the mandatory approvals for the provisional results. Refusing to accept this reason, the Standing Committee on Finance, on July 28, 2022, noted:

“The Committee are constrained to note that the Ministry has furnished a routine reply regarding the growing delay in the release of the Economic Census (EC) making the data on some of the items/samples **irrelevant or outdated** for stakeholders to be used as parameters. They have now sought to shift the onus on the [S]tates/UTs for not being forthcoming in giving approval of provisional results of the 7th EC. The Committee are, however, of the view that it is the responsibility of the Ministry to follow up with the States/UT for early examination of provisional results by the State level coordination committees to enable Ministry in releasing all-India results **at the earliest**. Thus the Ministry has to **squarely assume responsibility** in this regard as it is being conducted by them as a central sector sub scheme. It is high time that the Ministry should expedite the process of enumeration and produce reliable data publicly. The Committee desire that the identification of [S]tates should be done which are lagging behind in giving their approvals and accordingly corrective measures to be taken to **identify the problem areas** and difficulties encountered by these [S]tates and appraise the Committee in this regard.”⁴⁰ [Emphases by author.]

With most administrative activities moving online, sourcing such data has now become easier. In contrast, direct ground level data collection through surveys and censuses have become extremely problematic with dwindling cooperation from respondents, access restrictions, and data quality issues arising from a lack of trust in the ultimate objective of such data collection by official agencies. However, these factors are not to be seen as unsurmountable hurdles given that a large number of censuses and

⁴⁰ Lok Sabha Secretariat. 2022. *Standing Committee on Finance (2021-22)*, Seventeenth Lok Sabha, Ministry of Statistics and Programme Implementation, Fifty First Report, August, p.11. [http://164.100.47.193/lsscommittee/Finance/17_Finance_51.pdf].

surveys continue to be done by government agencies. The unresolved issue remains the reluctance on the part of governments to open up their administrative data bases for research.

IX. THE WAY FORWARD

“Statistics appeals to our rational side, our hearts providing a balance to our sometimes wayward hearts. In the data rich world that is emerging as our future, those nations, governments, businesses and individuals who use the power of numbers will prosper. Those who ‘get stats’ will get on. Those who do not will get left behind.”

John Pullinger

Presidential Address to The Royal Statistical Society, 2013⁴¹

An already strained Indian statistical system now faces a crisis of credibility as the autonomy of statistical agencies become visibly subordinate to political expediency. This is a major cause for concern, not only for the immediate implications but also for the long-term consequences of both data integrity and correctly charting the nation’s progress.

A case in point is the decennial population census that in normal times would have been held in 2021 but has not yet commenced due to the COVID-19 pandemic. The delay was understandable; but not its indefinite postponement. One can feel the absence of current census population figures has on local level planning, with most local bodies either using outdated 2011 data or making some crude projections. Even before the pandemic, the Census process for 2022 was in the news with the Government tagging the preparation of the National Population Register (NPR) and the preparation of the National Register of Citizens (NRC) based on the data collected during the census operations. Elsewhere, I have struck a note of caution on how this will have serious ramifications on the quality of census data and the need to detach it from an already stretched-out census process⁴². The Union government’s position is that the Census will be a “dynamic” one after proposed amendments to the Registration of Birth and Death Act, 1969,⁴³ but there is no indication of the Census of India, which has reportedly been postponed “until further orders”,⁴⁴ commencing anytime soon. This does not auger well for India’s statistical architecture that is already an international laggard. The decision to put off the Census “until further orders” also makes citizens even less aware of the state of affairs in India, in numerical terms, which the decennial exercises were providing. This would only go on to create a less-informed citizenry, vulnerable to claims

⁴¹ Pullinger, J. 2013. Statistics making an impact. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, October. Vol. 176. No. 4. Pp.818-839 [https://www.jstor.org/stable/43965360].

⁴² Mohanan, P.C. 2019. *NPR: A statistical nightmare*, Financial Express, December 31. [https://www.financialexpress.com/opinion/npr-a-statistical-nightmare/1807957/].

⁴³ Singh, V. 2021. *COVID-19 curbs off, but Census still on slow burner*, The Hindu, July 16. [https://www.thehindu.com/news/national/covid-19-curbs-off-but-census-still-on-slow-burner/article65647312.ece].

⁴⁴ Hindustan Times. 2022. *Census 2021 deferred until further orders: Centre*, July 27. [https://www.hindustantimes.com/india-news/census-2021-deferred-till-further-orders-centre-101658861900488.html].

that cannot be substantiated with any degree of credibility of authenticity. This, in turn will deeply erode the central space occupied by India's official data agencies.

The primacy of space occupied by Indian statistical agencies to collect and disseminate data in conformity with international commitments and scrutiny needs to be restored. While political use of data is unavoidable and the reticence of official agencies in entering such debates is understandable, the remedy lies in transparent dissemination of data with full meta-data.

The need to revitalise state statistical systems

Given India's size and complexities, large scale data gathering exercises are costly. A defining moment in the development of India's official data system was when Mahalanobis successfully articulated the need for scientific sample surveys as a cost effective solution. Underlying this was the careful planning and execution that went behind the process. One has only to peruse the correspondences between Prof. Mahalanobis and Prof. Dandekar⁴⁵ to understand the meticulous care these two stalwarts took in framing the questions and methodology in the first national sample survey. Unfortunately, we do not find such careful designing of surveys now. An example for this is the MoSPI withholding the results of 2017-18 household consumer expenditure survey citing data quality issues. The consumer expenditure survey is one survey that got worldwide recognition and became the flagship Living Standard Measurement Surveys actively recommended by World Bank. For such an important source of information, which has direct implications on the nation's economy, the people are still in the dark as to what the "quality issues" are that made it necessary for the government to withhold its release.

While it is fashionable to ask for censuses of all kinds, socio-economic changes can be better understood thorough careful sample studies. Over the years, most major statistical surveys are outsourced to private agencies or done with the help of contract workers slowly dissipating the expertise and legacy built on solid scientific foundations.

India is also a pioneer in the dissemination of micro-data. Moreover, a large number of data archives have been built within India and elsewhere, through which researchers can now access these data. While anonymisation is rather easy for survey data, large administrative and census based data requires careful curating to retain respondent confidentiality. In this context, the proposed data protection statute needs

⁴⁵ Dandekar, V.M. 1953. *Report on the Poona Schedules of the National Sample Survey (1950-51)*. Gokhale Institute of Politics and Economics, Publication No. 26, Paperback, January 1. [<https://dspace.gipe.ac.in/xmlui/handle/10973/13909>].

to balance research needs an individual privacy of data. All these point to the need for more resources and specialization for the statistical agencies.

Revitalising the state statistical systems is also key to strengthening the federal system. Most of the state statistical systems have not kept pace with the changing times and function with poor infrastructure and severe budgetary constraints. Recently the governments of Kerala and Madhya Pradesh constituted State-level Statistical Commissions that are expected to address the issues. The centrally sponsored and funded operations like the Livestock Census, Agricultural Census, Economic Census, Irrigation Census etc. are usually done by State agencies. But these data remain with the central agencies and are not put to careful, analysis by the State agencies either due to lack of expertise or demand. In some cases the national level planning for these exercises also means that specific State-level issues are not addressed reducing their utility for the concerned State.

Proposed draft Data Governance Framework

The growing digitisation of economic and social intercourses among economic units and in access and delivery of public services has resulted in expansion of the data horizon like never before. This expansion is producing a diverse variety of data, with ever increasing velocity and volume of data generation. Unlike in the past, not all of this is in the hands of Government; rather a substantial part of this new data is generated by private agencies and at times with no centralised agency in control. In this set up, the question of data protection and privacy has therefore acquired urgency. Many national governments have come out with legal structures to manage, control and protect privacy of data as the unrestricted use of such data could be detrimental to national and personal interests.

The Government of India has also been grappling with myriad issues connected with data protection and had tabled a Bill in parliament. Statistical agencies need impersonal data but with as much details as possible keeping in view the wider interest in socio-economic changes. Conventional government thinking not only gave primacy to individual organisations but also assumed that they would own whatever data collected by them. It is only in recent times that governments are urged to think along the lines that provision of data for public use is considered their solemn duty.

We have already mentioned the efforts through the National Data Access and Sharing Policy (NDSAP 2012) that mostly dealt with data owned by Government agencies. Recently the Government has also come out with a Draft National Data Governance Framework Policy. This draft has a much wider

connotation of data, in particular the kinds of data emanating from the digital platforms that are noted as ‘empowering citizens, enhancing government-citizen engagement, and driving data-driven governance’.

The policy rightly admits that the digital government data is currently managed stored and accessed in a non-optimal way and proposes a new National Data Governance Framework Policy (NDGFP) that aims to realize the full potential of Digital Government.

This policy proposes a new office responsible for framing, managing and periodically reviewing and revising the policy called the “India Data Management Office (IDMO)” under Ministry of Electronics and IT. Each Ministry/Department shall have separate Data Management Units headed by a designated data officer. State Governments are also to be encouraged to designate/appoint State Level Data Officers. The IDMO shall coordinate closely with line Ministries, State Governments, and other schematic programs to standardize data management by building up capacity and capabilities in each Ministry. Further it will accelerate inclusion of non-personal datasets housed with ministries and private companies into the India Datasets programme. It has among other objectives the standardisation of data management and security standards across the whole of Government, creation of common standard based public digital platforms. For purposes of safety and trust, any non-personal data sharing by any entity can be only via platforms designated and authorised by IDMO. It will also build a platform that will allow dataset requests to be received and processed. The proposed IDMO will have say in data storage and retention, Government-to-Government Data Access, and many other standard data management activities cutting across all agencies and domains.

The proposed data governance framework is a clear indication of the deficiencies in the present system in deriving the full potential of the new information technology based architecture in various government programs and processes. This should also be seen as a part of the historical evolution of data management in the government starting with the post-independence statistical architecture based on surveys, censuses and administrative data and the setting up of the first Government Computer Centre in the Ministry of Statistics to cater to the computing needs of the Government. Subsequently, the National informatics Centre (NIC) was set up to help Ministries to computerise their activities and develop and host data bases. Later we had the NDSAP under the aegis of the Department of Science and Technology and the Open Government Data portal supported by NIC. Efforts were also made to develop a National Spatial Data Infrastructure (NSDP) to bring all spatial data with common meta-data

standards under one platform. The need for a common meta-data standard for all e-governance projects was also recognized through the development of a separate portal for this.

There is no doubt that the evolving 'platformisation strategy for delivery of public services and governance offers, immense scope for data based research and increased efficiency in governance. The data generated would also make significant contributions to the official statistical system. The basic premise on which official statistical system is built has been that such statistics should represent the whole country and its population or at least cover clearly defined domains. It is this that makes survey agencies take care to include all groups and geographical territories so that the estimates truly represent the whole country and its economy. The inclusivity of the data generated from the new strategy would therefore depend on the inclusivity of the process. Any systematic exclusion would clearly bias the conclusions derived from such data. Does this all pervasive digitalisation exclude economically and socially marginalised people?

Another issue of concern to statistical agencies is the need to produce data consistently and on a continuing basis. It is well known that the administrative processes are susceptible to changes and that such changes are not always linked to the past processes. Statistical agencies would find it difficult to use these data sources if their coverage or definitions get updated or changed. One of the reasons the statistical agencies have not been able to make use of the MIS of various Ministries and their programs could be attributed this.

Given the compartmentalisation of subjects dealt by Ministries both at the centre and States, one cannot be too optimistic about the proposed framework bringing perceptible improvements in the statistical system. The statistical system depends on a variety of concepts adopted from different fields for generating indicators; be it the macroeconomic aggregates or demographic parameters or any other social or developmental indicators. Domain knowledge has to be the basis for the underlying design in the proposed integration of platforms.

Statistical literacy and role of official statistical agencies

Before concluding, a short note on the importance of statistical literacy will be in order. This is more so given the importance of a general understanding of statistics by the citizenry. There is an oft repeated "paraphrased" quote attributed to H.G. Wells that says, 'statistical thinking will one day be as necessary

for efficient citizenship as the ability to read and write^{46,47,48}. Though this observation was made in a Presidential Address to the American Statistical Association in 1950, it finds renewed relevance when data has become all pervading. With the advancement in general literacy and the advent of social media an individual is flooded with information, or rather processed data from diverse sources with presented and interpreted to suit specific purposes. A certain amount of statistical literacy would be a real necessity to assess the claims and counter claims on issues affecting the public.

Gunnar Myrdal, the Swedish economist in his work on South Asian countries had noted that the ability not only to read and write figures with understanding but also to add, subtract, multiply and divide is of importance in economic and social development. In the current context one can safely add that certain basic statistical literacy is also essential, at least for those joining public debates. While there is no way to measure this, not to talk of enforcing any of it, the national statistical agencies have a duty to explain official numbers and provide the right perspective. A case in point is the press release issued by the Ministry of Health and Family Welfare to refute the WHO estimates of COVID-19 mortality for India⁴⁹. The Ministry on behalf of the Government strongly objected to the use of mathematical models for projecting excess mortality estimates in view of the availability of authentic data and said that the validity and robustness of the models used and methodology of data collection are questionable. While there is no doubt that the Union Ministry of Health and Family Welfare has adequate statistical resources, such a technical matter relating to national data should have been left to the NSO to deal with rather than giving a political colour to the statistical method used by WHO by the involvement of the concerned Ministry.

Explaining numbers is not always very easy and is not usually part of the statistical curriculum in universities. While statistical communication has not been a very strong point for government statisticians, a silver lining is the evolution of data journalism. Major newspapers are now devoting space for explaining data, particularly those from government sources, by a new generation of qualified data journalists. The increased public interest in data is also seen in the popularity of books on data. The importance of statistical literacy is reflected in the Rangarajan Committee Report in which mentions

⁴⁶ **Wilks, S.S. 1951.** [Undergraduate Statistical Education](https://www.jstor.org/stable/2280089), *Journal of the American Statistical Association*, March 1951, Vol 46, No. 253, P.5. [https://www.jstor.org/stable/2280089].

⁴⁷ The Consortium for the Advancement of Undergraduate Statistics Education (CAUSE), points out that Wilks was “paraphrasing” what H.G. Wells wrote in his 1903-book, [Mankind in the Making](https://www.gutenberg.org/files/7058/7058-h/7058-h.htm). [https://www.causeweb.org/cause/resources/library/r1266].

⁴⁸ **Wells, H.G. 2003.** [Mankind in the Making](https://www.gutenberg.org/files/7058/7058-h/7058-h.htm), The Project Gutenberg, March 3. [https://www.gutenberg.org/files/7058/7058-h/7058-h.htm].

⁴⁹ **Ministry of Health and Family Welfare. 2022.** [Excess Mortality Estimates by WHO – India Strongly objects to the use of mathematical models for projecting excess mortality estimates in view of the availability of authentic data](https://pib.gov.in/PressReleasePage.aspx?PRID=1823012), [Press Release] *Press Information Bureau*, May 05. [https://pib.gov.in/PressReleasePage.aspx?PRID=1823012].

“increasing and promoting public awareness of official statistics” as one of the functions of the NSC in order to “improve public trust in statistics”⁵⁰

In summary, the Indian statistical system faces many challenges; none of which are insurmountable. There is a confluence of many developments which create a favourable environment to mainstream statistics in not only policy making but also shaping informed public decisions: technology for data acquisition and processing are available like never before, interest in data is at its highest, access to data through the internet is more affordable and nearly ubiquitous. What is wanting, however, is the wider political awareness about the importance of data among India’s citizens. Institutional mechanisms that insulate the country’s data collection, analysis, release, and dissemination processes are an important requirement.

⁵⁰ Ministry of Statistics and Programme Implementation. 2001. op. cit. p. 83.



About the Author

P.C. Mohanan is former Acting Chairman, National Statistical Commission (NSC), He was earlier a Member of the National Statistical Commission from June 2017, and the Acting Chairman of the Commission from October 2018 until his resignation from the position in January 2019. He entered the Indian Statistical Service ranked Second in the 1979 batch and worked in both the NSSO and the CSO until his retirement in 2015. He has been a member of important technical committees that have addressed issues in India's socio-economic sectors, and has held international consultancy assignments in the Asian Development Bank, UNDP, Food and Agricultural Organisation, and International Labour Organisation. He can be contacted at parachottilmohanan@gmail.com



THE HINDU CENTRE

for

Politics and Public Policy