



THEIR VIEW

# Measuring India's manufacturing sector remains a data challenge

*Despite revisions of data sources and methodologies, this statistical exercise is full of complexities that are hard to untangle*



**BIBEK DEBROY, AMEY SAPRE & ADITYA SINHA**

are, respectively, chairman, Economic Advisory Council to the Prime Minister (EAC-PM), associate professor, National Institute of Public Finance and Policy (NIPFP) and additional private secretary (policy and research), EAC-PM

In our previous column ([bit.ly/3NTggMg](https://bit.ly/3NTggMg)), we brought out some of the critical conceptual and paradigmatic shifts of the 2011-12 data series, which marked a significant departure from the previous series. One of the intensely discussed changes was the introduction of the MCA21 data-set for the private corporate sector, which includes organized parts of both the manufacturing and services sectors.

The MCA21 data-set was introduced in place of the Annual Survey of Industries (ASI) to expand coverage of the organized sector, as opposed to factories in case of ASI. Alongside manufacturing, the data-set also captured service-sector companies, which significantly improved coverage compared to the previous series.

From a systemic perspective, it is undeniable that a large, complex and fast-moving economy requires quality data-sets and sound measurement techniques for macro-aggregates. However, the experience of the 2011-12 series has shown mixed results. The MCA21 data has almost entirely overshadowed the debates, even though an equal number of problems existed in measuring unorganized parts of the Indian economy.

Even within the confines of MCA21, the benefits were accompanied by several complications, many of which persist as contentious issues, difficult to mitigate. In comparison with the deficiencies of the ASI, the MCA gave rise to significant measurement challenges, many of which have been illuminated in published research and reports from the National Statistical Commission.

The MCA21 represents an administrative data-set, fundamentally distinct in structure from the ASI. The deployment of such data-sets for national accounting purposes is a complex endeavour, necessitating extensive coordination with ministries for comprehending processes of data collection, definitions, objectives and eventual applications. These data sets are not collected or generated via any statistical design, schedule or questionnaire intended for statistical purposes. Their primary function is to facilitate regulatory monitoring and mandate compliances on an array of legal statutes. So the statistical agency finds itself wrestling with data structures that emanate from regulatory practices, accounting standards and the administrative processes of the ministry responsible for data generation, over which it has no control. Thus, the transition from ASI to MCA involved replacing survey data with administrative data, ushering in many attendant issues.

Technical details aside, the present compilation process of basic aggregates has major limitations. Compared to ASI, the MCA21 data-set has both manufacturing and service-sector companies.



BLOOMBERG

MCA21 lacks clear identifiers of economic activity within the registered entities for correctly classifying companies into respective sectors. Enterprises are categorized based on the NIC digits present in the Corporate Identification Number, which may not necessarily align with actual economic activity conducted by an enterprise. Limited information on product-level revenue is available in other forms, but also has considerable limitations. The problem is compounded in case of diversified enterprises that have multiple products and services and operate at several locations. Misclassification distorts sector-wise estimates, and such estimates show a disconnect with other measures of industrial activity.

The data also lacks geographical indicators, making it problematic for computing state-level aggregates. State-level aggregates for the organized part of the manufacturing sector are allocated from national totals to each state based on shares of respective economic activities in ASI. This method has the unrealistic assumption that while the structure of state-wise manufacturing is as shown by ASI, its value added is allocated from the national total, which is based on MCA.

As of January 2023, the MCA21 data offers close to 2.4 million companies, of which nearly 1.5 million (62%) are considered active. Their economic activity classification is as per NIC 2004 (instead of NIC 2008) and the data shows 21% of active companies as 'manufacturing'. The universe of companies is dynamic in the sense that the MCA adds companies on a monthly basis (based on new registrations) and de-registers companies as per the norm of de-registration. For this administrative data to be robust for national accounts purposes, it is necessary that filing patterns be studied in depth. In the 2019 press note on GDP ([bit.ly/3Bkl-gYg](https://bit.ly/3Bkl-gYg)), data from 2012-17 showed that on average, about 60% of active companies file their financial statements and are thus available for estimation.

This landscape of companies is considerably different from the universe captured by ASI. Unless a year-on-year mapping of ASI and MCA is achieved, the material gains from the shift remain unclear. Unlike ASI, it is also impossible for the statistical agency to physically verify enterprises, unless an enterprise sampling is undertaken using the list of registered companies. The experience of the 74th Round NSS Survey on Service Sector Enterprises had essential findings on this front.

As it stands, at best, MCA21 data offers a broad aggregate picture of our corporate sector if we concede the classification problem. It offers wider coverage if we concede another on quality of data and the universe of companies. It offers broader scope to estimate value addition if we concede yet another on a distinction between manufacturing and services. It also offers a faster way to prepare annual estimates, if we concede issues of a time lag of ASI and the need for following a bottom-up estimation approach for state GDP. These limitations do not automatically make the ASI preferable, but leave more questions than answers.

The influence of the MCA21 data was not confined to organized sectors. A substantial hurdle persists in quantifying the unorganized portion of manufacturing. As it stands, estimates lean heavily on the 2010-11 survey (67th Round of Unincorporated Enterprises) and directories of urban and rural establishments. These resources, however, have waned in relevance due to the profound transformations in the enterprise landscape, particularly after GST implementation and the cataclysmic effects of covid. Moreover, GVA levels are extrapolated by applying growth rates of the organized sector from the MCA, which potentially leads to significant measurement inaccuracies. In its current state, assessing India's manufacturing sector poses a formidable challenge, and despite revisions in data and methodologies, it has added complexities that are arduous to untangle.