

On the pitfalls of estimating GDP

The ongoing GDP series is due for revision. The official agency is considering using GST data to estimate value addition. However, the agency should be careful to not apply unverified datasets and shaky methodologies without adequate testing and validations for GDP estimation

ECONOMIC NOTES

R. Nagaraj

Gross Domestic Product, or GDP, is the most significant measure of a country's economic size. It is also a universal denominator for comparing indicators across countries and regions or for sizing up tax burdens or welfare expenditures. GDP is usually more meaningful at "constant" prices or in "real" terms – netting out the effect of price changes. The real GDP is estimated for the "base year", requiring a variety of datasets on output, prices, and employment. Every 5-10 years, the GDP base year is revised to account for changes in relative prices and output composition. The National Statistical Office (NSO) is tasked with "revising" the GDP series, usually drawing upon expertise from many fields.

The ongoing GDP series with the base year 2011-12 is due for revision. 2020-21 is the proposed new base year. All required major datasets are said to be available except for Census data. The NSO is considering using the goods and services tax (GST) data to estimate value addition, replacing the currently used Ministry of Corporate Affairs' MCA-21 database for the Private Corporate Sector (PCS), which accounts for about 38% of GDP.

Why the change?

After all, the MCA-21 database was brought in only in the last revision, with 2011-12 as the base year. Previous to that, the Annual Survey of Industries (ASI) was the long-standing workhorse for estimating factory manufacturing value-added. The Reserve Bank of India's (RBI) small sample of large companies, with the majority paid-up capital of PCS, was used to estimate the non-financial corporate sector output. The statistical agency changed it to the MCA-21 database as the ASI claimed to miss out on value addition outside of factory premises in a corporate entity. Likewise, reportedly, the



GETTY IMAGES

RBI sample was inadequate to account for the rapidly growing PCS. Moreover, the availability of the extensive and up-to-date MCA-21 data, obtained from the mandatory filing of corporate annual returns and quarterly corporate results – it was contended – would enable fuller capturing of the corporate output.

The 2011-12 base year GDP (replacing the 2004-05 base year series) showed a marginally smaller absolute GDP size and a faster growth rate. But for the manufacturing sector in 2013-14 at constant prices, the annual growth rate was (+) 5.4% in the new series, compared to (-) 1.90% in the earlier series. Such a sharp divergence in the rate and direction of industrial growth by the two GDP series was a surprise. Moreover, the upward revision of the industrial growth rate didn't square with related macro aggregates, such as bank credit growth or industrial capacity utilisation, leading to

widespread scepticism of the new GDP estimates. Statistical investigations zeroed in on an untested or inadequately vetted MCA database as the source of the overestimation problem.

The official agency, however, defended its new estimates, claiming they capture value addition more completely, using a much more extensive database, improved estimation methods, and following the latest template of international best practices. Critics, however, wondered if a bigger dataset is necessarily a better data set. And if the new estimates were better or overestimates. The statistical dispute remained unresolved as the government refused to make the MCA data available for independent scrutiny or reveal its estimation methodology for verification.

Systematic overestimation

With time, however, it has been possible to compare estimates of Gross Value

Added (GVA) in the manufacturing sector as per GDP series (in the National Accounts Statistics) and by the ASI – based on production accounts of registered factories – for a reasonably long period. We compared (i) GVA and (ii) Gross Fixed Capital Formation (GFCF) (fixed investment) at constant prices for 2012-13 to 2019-20 as reported by the NAS and ASI. The results were startling. The average annual growth rate of GVA in NAS was 6.2%, while it was only 3.2% in ASI. The difference was much sharper in GFCF: 4.5% by NAS and 0.3% by ASI, respectively. These comparisons show a systematic overestimation in NAS estimates (based on the MCA-21 database) compared to the ASI-based estimates, vindicating the doubts raised about the integrity of the GDP estimates.

The evidence presented here is a cautionary tale for the proposed use of GST data for GDP estimation. It's a stark reminder of the need for the official agency to guard against the hasty application of unverified datasets and shaky methodologies without adequate testing and validations for GDP estimation. NSO must initiate pilot studies to verify the GST dataset's suitability for value addition estimation of specific industries, sectors, and States. Such validation is crucial to ensure the estimation's truthfulness and instil

confidence in the integrity of the GST data. Alternatively, NSO could explore reverting to ASI to estimate GDP manufacturing, as the database is now available with a shorter time lag.

GST data can be a game-changer for GDP estimation in the proposed revision. It is a large and up-to-date database, however, its details are in a black box, as it has not been open for policy research. Without systematic analyses and cross-validation disaggregated by production and institutional sectors and regions by independent agencies, the validity of GDP estimates on GST data will be hard to establish.

R. Nagaraj is with the Centre for Liberal Education, IIT Bombay.

THE GIST

▼ Gross Domestic Product, or GDP, is the most significant measure of a country's economic size. It is also a universal denominator for comparing indicators across countries and regions or for sizing up tax burdens or welfare expenditures.

▼ For the manufacturing sector in 2013-14 at constant prices, the annual growth rate was (+) 5.4% in the new series, compared to (-) 1.90% in the earlier series. Such a sharp divergence in the rate and direction of industrial growth¹ by the two GDP series was a surprise.

▼ The evidence presented here is a cautionary tale for the proposed use of GST data for GDP estimation.