

## Analysis of revisions in Indian GDP data

AMEY SAPRE AND RAJESHWARI SENGUPTA, *Consultant, National Institute of Public Finance and Policy, New Delhi, India*

Compilation of GDP is a complex exercise that requires a combination of estimation methods and vast amounts of data from several sectors of the economy. Since collection of detailed sectoral data is time consuming, GDP estimates for any given year are released in a sequence of revised estimates depending on different levels of data availability. In India, the Central Statistics Office (CSO) releases six estimates of annual GDP for any given year, over a period of two years and 10 months. The initial estimates (or projections) are called the Advance Estimates (AE) and Provisional Estimates (PE). Over time, the initial estimates are revised and are released as the First (1st RE), Second (2nd RE) and Third (3rd RE) Revised Estimates.

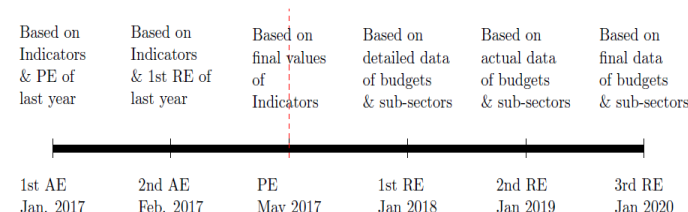
Typically, for various stakeholders and policy formulations, the initial GDP estimates are more relevant as they are available within the financial year. However, the true picture of the economy, in terms of the magnitude and direction of growth unfolds over the entire revision cycle. It is thus important to study the revision cycle in order to obtain a detailed picture of growth in the economy. In a recent [paper](#) we use historic data on revisions to answer two questions: (i) what has been the magnitude of revisions in GDP estimates at the aggregate and sectoral levels? and (ii) how can we gain predictability over revisions for each sector? In order to assess predictability, we develop a method to construct a confidence band around the AE using standard deviation (SD) of revisions. The analysis also covers; (i) discussion on short and long term revisions, (ii) high frequency indicators used for initial estimates and (iii) identification of sectors that have large and unpredictable revisions in growth rates.

### Process of GDP data revisions

In the revision cycle, the AE uses 7-8 months of data on high-frequency indicators that are assumed to capture the level of economic activity in each sector. The PE is the first full-year estimate as it is based on 12 months of data on all sectoral indicators. In this context, scholars have argued that data revisions contain both 'news' and

'noise' about the economy's growth performance (see, for instance, Mankiw and Shapiro (1986), Fixler *et al.* (2008)). As initial estimates are compiled with incomplete data or proxies based on high-frequency indicators, there is likely to be more noise in these estimates. Gradually with data availability, the extent of noise is expected to diminish, and the revised estimates start to reflect 'news' about the state of the economy. In the Indian context, this process can be visualised in the following figure.

**Figure1: Timeline of GDP data revisions: From noise to news**



Source: Sapre and Sengupta (2017)

### What do we know from studying revisions?

The magnitude of revisions in annual GDP growth rates from the initial to the final estimate has, in general, been low. However, the direction in revisions suggests that for most years, the AE underestimated the actual GDP growth rate. The figures also suggest that for most years the 2<sup>nd</sup> and 3<sup>rd</sup> RE have been close and consistently in one direction, implying that the focus of improving data needs to be on the AE as the discrepancy is much larger in these initial estimates.

Following international practice, we divide the revisions into short term (AE–1<sup>st</sup> RE) and long term (AE–2<sup>nd</sup> RE). We find that for manufacturing, mining, trade, and community services, the short-term revisions on average, have been negative, implying that the initial projections of the sectors were underestimates. In contrast, comparing long-term revisions (AE–2<sup>nd</sup> RE), we

find a reverse phenomenon for sectors such as electricity, construction, and financial services. Agriculture and trade are the only sectors where the short-term and long-term revisions have been in the same direction. Among all sectors, mining and quarrying shows the highest variability in long-term revisions. We construct confidence [bands](#) for each sector as it allows us to obtain a range around the AE. The band is constructed using 1 SD of revisions to compute upper and lower bounds around the AE, i.e.  $(AE - \sigma_{AE}, AE + \sigma_{AE})$ . To assess predictability, we count the number of times the final estimate falls within the confidence band for each sector. We find that the confidence band is accurate for all years for the electricity, gas, and water supply sector, while the accuracy is 70% in case of agriculture, construction, trade, financial services, and community services sector. The predictability is lowest at 57% for the manufacturing, and mining and quarrying sectors.

## Conclusion

Revisions in GDP data have important implications for policy formulations and business decisions. While Advance Estimates of GDP may be important for macroeconomic assessment, understanding the revision cycle helps us to distinguish between data changes and genuine macroeconomic fluctuations. However, revisions have their own share of problems. Large and unpredictable revisions raise concerns over data reliability. In the context of revisions in Indian GDP, historic revisions suggest that the focus ought to be on data quality and choice of high frequency indicators used for the initial estimates so as to contain the magnitude of revisions and improve accuracy over the revision cycle.

## Further Reading

Fixler, Dennis J and Bruce T Grimm, 2008. [The Reliability of the GDP and GDI Estimates](#), *Survey of Current Business*, February, 2008.

Mankiw, N. Gregory and Matthew D Shapiro, 1986. [News or Noise: An Analysis of GNP Revisions](#), *Survey of Current Business*, May (66): pp. 20-25.

Sapre, Amey and Rajeswari Sengupta (2017) [An analysis of revisions in Indian GDP data](#), National Institute of Public Finance and Policy Working Paper, No. 213, November, 2017.