



ILLUSTRATION: ROHITH PHORE

**LEKHA
CHAKRABORTY**

The author is professor, NIPFP, New Delhi



● DATA FUNDAMENTALISM

The rise of data self-censorship thresholds

The lack of adequate frequency in the release of Consumer Expenditure Survey data can spill over to the base year revisions of GDP and in also measuring inflation. The impact of public policies like demonetisation and GST on the informal sector cannot be analysed unless we have authentic expenditure data

THOMAS PIKETTY'S magnum opus 'Capital in the Twenty-First Century' and his later works popularised the notion of the 'top 1% income quintile' and estimated income inequality in the context of India. He used a combination of income tax data and the Consumer Expenditure Survey (CES) of the national sample survey rounds data (NSSO) to arrive at these estimates and triggered a debate on the rising income inequality in India.

The income tax data was released in October 2019 by the Central Board of Direct Taxes (CBDT). However, the Ministry of Statistics and Programme Implementation (MoSPI) has decided "not to release the CES results of 2017-18" in view of "data quality issues" in a statement. It was also reported that the ministry is "separately examining the feasibility of conducting the next CES in 2020-21 and 2021-22 after incorporating all data quality refinements in the survey process." The income inequality

estimation is just one exercise of many such empirical analysis using this CES data. As Somesh Jha of *Business Standard* tweeted, "This will be the first time in India's history that an official survey report will be scrapped without it being made public."

Researchers and analysts have been patiently waiting for the release of the 2019 National Sample Survey Consumption rounds—the CES data—to get the marginal per capita consumption expenditure (MPCE) quintile data to plausibly update these estimates and examine the extent of inequalities in India. Despite its methodological limitations, this analysis was all the more compelling at this point in time, against the backdrop of 'unintended consequences' of public policies on economic growth—for instance, demonetisation and the nebulous new GST regime in India. The Union Budget 2020 also hardly had any effective fiscal announcements to trigger the economy from the enormous glut it is into.

This decision by the government to scrap the data before releasing it has raised two concerns. One is the 'data fundamentalism' and two is the lack of high frequency data or the quinquennial surveys on crucial macro variables to analyse the economy, and the time lag involved in publishing the data.

The data fundamentalism relates to the suppression of data over 'adverse' inferences or inferior data. The 'leaked' NSSO consumption round report found that, in 2017-18,

consumer spending fell for the first time in more than four decades. As per the 'leaked' latest consumption expenditure survey by the National Statistical Office (NSO), the average amount of money spent by a person in a month fell by 3.7%, from ₹1,501 in 2011-12 to ₹1,446 in 2017-18 in 'real' terms. Technically, 'real' expenditure is nominal expenditure minus inflation. There is an 8.8% decline in rural MPCE (monthly per capita expenditure) and a 2% decline in urban MPCE. This is serious as it has clear linkages to the adverse impacts of demonetisation and GST, but the government is still on the denial mode.

Martin Ravallion, an eminent macro policy economist, has tweeted that, "If

verified when the data are released, there are at least two plausible explanations for this alarming reversal for India's poor: the demonetisation and the introduction of the GST. Both transient shocks, but costly."

Ex post to demonetisation, India has not conducted any impact analysis of demonetisation for the lack of authentic data. The data on 'household income' is not available in India, so researchers have been using MPCE of households as a proxy for income in the analysis of inequalities. It is interesting to recall here a study on the impact of demonetisation done by IMF Chief Economist Gita Gopinath and her team using satellite data of 'nightlight intensity' or 'luminosity' as a proxy for income. However, there are constraints in interpreting the inferences drawn from the satellite data.

Now with the withdrawal of the present NSO CES rounds, we have to wait at least till 2022 to get the MPCE data, and it will be too late to analyse the effectiveness of policies like demonetisation and GST and to frame adequate impulse response functions to the public policy failures, if any, or to the extent of its impacts.

The leaked NSO data also revealed a fall in the consumption of food items, which has given a clear indication that malnutrition can yet again be a silent emergency in India. If we co-read this data with the estimates from recent first-ever Comprehensive National Nutrition Survey (CNN) of India 2019, the picture will be clearer. The anthropometric data analysis from the recent comprehensive nutrition survey revealed that, in India, 35% of under-5 children are stunted, 17% are wasted, 33% are underweight. "The triple burden of malnutrition—undernutrition, hidden hunger and overweight—threatens the survival, growth and development of children, young people, economies and nations," highlighted the report published by the health ministry in coordination with the UNICEF.

Is this the first time 'data fundamentalism' due to adverse inference happened in India? The answer to this question is not in the affirmative. The data on back series of GDP calculated with the revised 'base year' was removed from the website of MoSPI a year ago. The reason, officially mentioned, was the deficiency in methodological advancements and comparative data paucity.

Yet another instance was the controversy over the Periodic Labour Force Survey (PLFS) for 2017-18. The 'leaked' PLFS inferences revealed the unemployment rate in the country to be 6.1%, a four-decade high. However, the government released both the report and the unit-level data subsequently and it is uploaded in the MoSPI website this year.

The lack of adequate frequency in the release of CES data can spill over into the base year revisions of GDP and in also measuring inflation. The impact of public policies like demonetisation and GST on the informal sector cannot be analysed unless we have authentic CES data.

The integrity of the national statistical system is crucial for designing, commissioning and correcting evidence-based public policy in India. Without data, anything said is just another 'opinion' and it cannot be an analytical input. And that is definitely not the way public policies are judiciously framed.

The silver-lining is the recent decision made by the government in constituting an expert committee led by former Chief Statistician Pronab Sen to conducting fresh surveys on consumer expenditure with revised methodology for FY 21 and FY 22. This survey will be crucial for receiving the estimates on poverty and inequality in India, even though one is not sure about the quality of responses against the backdrop of controversial CAA/NRC decisions. The 'self-censorship threshold' of people would definitely affect the quality of data generated.

By withdrawing the current NSO CES rounds, we have to wait till 2022 to get MPCE data; it'll be too late to analyse the effectiveness of policies like GST, demonetisation

