Consumption-based poverty estimates have relevance

C. Rangarajan
was Chairman, Expert Group to Review the Methodology for Measuring Poverty (2014)

S. Mahendra Dev
was Member, Expert Group to Review the Methodology for Measuring Poverty (2014)

A recent report by NITI Aayog on multidimensional poverty shows that the percentage of the poor has gone down from 25% in 2015-16 to 15% in 2019-21 and around 135 million people were lifted out of poverty during this period. The Global Multidimensional Poverty Index report of 2023 of the United Nations Development Programme (UNDP) and the Oxford Poverty & Human Development Initiative (OPHI), which was released recently, also shows that the incidence of the multidimensional poverty index declined from 27.5% in 2015-16 to 16.2% in 2019-21.

In this context, we briefly examine the issues, particularly on methodology relating to the multidimensional poverty index, and argue that consumption-based poverty estimates are still very relevant. Multidimensional poverty estimates are not substitutes for National Sample Survey (NSS) consumption-based poverty ratios. In the end, we also flag some concerns about consumption expenditure surveys and the need to correct them.

Comparison of results

The report of the Global Multidimensional Poverty Index (MPI) 2018 says: "India has made momentous progress in reducing multidimensional poverty. The incidence of multidimensional poverty was almost halved between 2005/06 and 2015/16, climbing down to 27.5 per cent. Thus, within ten years, the number of poor people in India fell by more than 271 million – a truly massive gain". This is high praise indeed.

Is the conclusion of global MPI a new revelation? No, as far as the 2015-16 estimates are concerned. The estimates of poverty based on consumer expenditure and using the Tendulkar committee methodology show (over a seven-year period between 2004-05 and 2011-12) that the number of poor came down by 137 million despite an increase in population. According to the Rangarajan Committee methodology, the decline between 2009-10 and 2011-12 is 92 million, which is 46 million per annum. For a decade, it will be larger than that of global MPI. However, in absolute terms, the poverty ratios based on the Tendulkar and Rangarajan Committee methodologies are lower than as estimated by global MPI.

The search for non-income dimensions of poverty possibly stems from a view that in terms of the capabilities approach to the concept and measurement of poverty, some of these 'capabilities' may not be tightly linked to the privately purchased consumption basket in terms of which the poverty lines are currently drawn. Therefore, poverty based on income or consumption is different from deprivations based on education or health.

As pointed out by the Expert Group to Review the Methodology for Measuring Poverty (2014), there are reservations on using multiple indicators as these multidimensional indicators/measures raise several issues regarding their measurability, aggregation across indicators, and, crucially, of databases that provide the requisite information at reasonably short intervals. These need to be considered and evaluated carefully. For example, there is a problem with the child mortality indicator as it is for population groups and not for households.

Aggregation is another problem. In principle, they should be independent. Access to safe drinking water, for example, cannot be aggregated with indicators such as child mortality. Even in respect of independent indicators, analytically appropriate rules of aggregation require that all of them relate to the same household. More generally, this requirement poses several data constraints.

It may be noted that we are not against multidimensional poverty or deprivations. One can analyse the progress of non-income indicators such as education, health, sanitation, drinking water, and child mortality over time with income or consumption poverty. But, converting all of them into an index poses several problems. Deaton and Drèze (2014) also indicate that "it is important to supplement expenditure-based poverty estimates with other indicators of living standards, relating for instance to nutrition, health, education and the quality of the environment".

On multidimensional issues, Srinivasan (2007) says viewing public services as another dimension besides consumption in a multidimensional conceptualisation of poverty is more fruitful. However, he is critical of multidimensional indices. He says that "collapsing many relevant but not necessarily commensurate dimensions into a single index defined as an arbitrarily weighted sum of disparate indexes makes little sense. The Human Development Index pioneered by the United Nations Development Programme is an example of an arbitrarily weighted sum of non-commensurate indexes. It certainly is not a multidimensional conceptualisation in any meaningful sense but simply yet another arbitrary unidimensional index".

In the minds of most people, being rich or poor is associated with levels of income. The various non-income indicators of poverty are in fact reflections of inadequate income. Defining poverty in terms of income or in the absence of such data in terms of expenditure seems most appropriate, and it is this method which is followed in most countries.

We do not have official data on consumer expenditure after 2011-12 to make a comparison with trends in the multidimensional poverty index. The survey data on consumption expenditures done in 2017-18 have not been released officially. In the absence of such data, there have been several studies on poverty using indirect methods and using Centre for Monitoring Indian Economy (CMIE) and Periodic Labour Force Survey (PLFS) data sources – and they have come up with differing conclusions.

Need for changes in surveys

The consumption expenditure survey is being conducted in the current year. For purposes of comparison, we need to follow one method. Therefore, it is best to wait for the survey results to be published. Earlier surveys clearly indicate that the poverty ratio comes down strongly during a period of high growth. If you look at recent years including the COVID-19 period, the growth rate has come down. There is ground to believe that the rate of reduction in the poverty ratio must have slowed down. This is at best a guess. We need to wait for consumption expenditure survey data.

An important issue is the differences in aggregate consumption estimates between National Accounts Statistics (NAS) and NSS data. These two estimates of consumption (NSS and NAS) do not match in any country; India is no exception. What is perplexing is that the difference in India between the NSS and the NAS consumption is widening over time. From a difference of less than 10% in the late 1970s, it has come to 53.1% in 2011-12, i.e., the Survey Estimate is only 46.9% of NAS estimates. The difference is too big to be brushed aside. The National Statistical Office must study the problem and come out with possible suggestions to improve the collection of data through both routes.

In addition, there is a need to supplement the results of consumption surveys with a study of the impact of public expenditure on health and education of different expenditure classes.

Multi-dimensional poverty estimates are not substitutes for National Sample Survey consumption-based poverty ratios

Understanding the 'dance steps of numbers'

In a write-up in *Sankhyā: The Indian Journal of Statistics*, in 1935, Tagore referred to the magic of mathematics as "the dance steps of numbers in the arena of time and space, which weave the *maya* of appearance, the incessant flow of changes that ever is and is not". As today's data-obsessed world aspires to make data-driven decisions on every bit of life and lifestyle, one has to determine how the "*maya*" is woven and what data might genuinely offer.

Analyses and findings

An example first. The findings of a study that provided independent analysts with nearly four million words from roughly 8,000 comments in an online academic forum were described in a paper in *Organisational Behaviour and Human Decision Processes* (2021) by a large collaborative group of scientists. The study posed two straightforward hypotheses about how gender and academic status influence contributions to the forum's discussions. Within 29 studies, radically varied analyses and dispersed empirical findings were reported; several cases had significant effects in opposite directions for the same research question. Had it been only a single analyst using a single method, one would have stopped at that result.

What is mentioned above is but one instance of conflicting conclusions drawn from the same data by several experts. Such situations may, of course, cause a great deal of controversy if the data and/or topic are of public interest. Such a situation may seem weird at first. As a result, the common person is perplexed; they are unsure about the causes of these discrepancies and, more crucially, whose analysis to trust. Even experts struggle to comprehend the merits of various contradictory conclusions.

Well, statistical analyses are not unique, especially when different methodologies are used. Interestingly, depending on the models and associated margins of error, all findings could be correct, even if they are contradictory in nature. However, certain analyses may contain significant errors. But it is never easy to identify them, even for experts. The history of civilisation shows that the conclusions drawn from the simple observation that the sun crosses the sky every day could even be varied and contradictory. It took centuries to settle the issue. Thus, domain expertise is necessary; generic statistical aptitude alone is insufficient, so far as data analyses are concerned.

Atanu Biswas

is Professor of Statistics, Indian Statistical Institute, Kolkata

Getting to know the causality underlying a phenomenon should serve as the foundation for any data analysis

Anchored to a 'model'

How is it possible that the same data could lead to contradictory conclusions? This is normal, though, as statistical analyses rely on an underlying "model" to explain the pattern of the data. The decision-making processes depend heavily on these models, the relevant metrics, and the associated margin of error for drawing conclusions. While any statistical decision is subject to some margin of error, a common person is frequently unaware of this. Additionally, models are created using the expertise and judgement of the statisticians concerned. Interestingly, George E.P. Box, a British statistician, is credited with the famous quote: "All models are wrong, but some are useful". And a "suitable" model is frequently not statistically chosen before doing an analysis.

Thus, understanding the causality underlying a phenomenon should serve as the foundation for any data analysis. Let us consider a well-known example: the gender bias data for admissions for the autumn of 1973 at the University of California, Berkeley. According to the data, just 35% of women applicants were admitted, compared to 44% of men who applied – a significant enough difference to be explained by chance. A comprehensive examination of the data, however, revealed that different departments had varying degrees of difficulty admitting students; and that men tended to apply to less competitive departments with higher acceptance rates, while women tended to apply to more competitive departments with lower admission rates. The department-wise data actually revealed a "small but statistically significant bias in favor of women".

In Delhi, in 1947

Sensing the heartbeats of data is a tricky task in itself. Another example. A few members of a minority community took refuge in the Red Fort during the communal riots in Delhi in 1947. The number was unknown to the government, and contractors hired to feed them charged high amounts. A team from the Indian Statistical Institute (ISI) obtained three estimates of the number by dividing the quantities of rice, pulses and salt used per day, as quoted by the contractors by the respective per capita requirements of these commodities. The estimate obtained by rice was the largest, and the salt estimate was the least. Since rice was the most expensive, its quantity was presumably inflated. And because salt was inexpensive, the contractors did not use it that way. The ISI team suggested using the quantity obtained from salt as an estimate of the number of refugees.

It demonstrates how statistical expertise, wisdom, and a solid understanding of reality are all essential for fully grasping "the dance steps of numbers". But extracting the best out of data is never easy. Only top statistical experts, along with experts with domain knowledge, may, at most, aspire to perform that job.