**[The power of data](https://www.dawn.com/news/1583318/the-power-of-data)**

[Hadia Majid](https://www.dawn.com/authors/7298/hadia-majid) 05 Oct 2020

IN many ways, Covid-19 has highlighted the need for data, both scientific and behavioural — the former so that we can map the virus while making strides towards a cure, the latter to better enforce and understand the economic and social effects of pandemic-related measures. But what does socioeconomic data look like, and what can it tell us? More importantly, what are its limitations?

The Ehsaas Emergency Cash Programme required identification of the most vulnerable sections of the population. The first order of business became categorising the poorest. For this, policymakers turned to the large, detailed database being maintained to target households for BISP. This database lists household assets and socio-demographic characteristics that allow for the creation of a poverty scorecard. Yet, having worked with this and other databases, I know all too well that there are both exclusion and inclusion errors in the targeting of BISP transfers. Not only are programme benefits not going to those who deserve them, in some cases, they are being received by those who are not poor.

It is essential to be mindful of the biases that may enter data.

Some may argue that these kinds of errors have more to do with the political economy of government programmes and less with the inadequacy of data. However, the literature recognises that despite decades worth of research on the determinants of poverty, we still can’t fully predict who will experience it and under what circumstances. This became evident during the pandemic when households possessing the assets and skill base to not be categorised as poor reported increased vulnerability and food insecurity because of the unprecedented economic fallout.

Here, occupations we associate with increasing resilience to poverty, eg those in the service industry, and jobs we believe increase vulnerability to economic hardship, eg those in agriculture, reversed their positions. Clearly, it is not enough to simply list the various features of households to come to grips with how they will fare under different circumstances. So, what should data look like?

Article continues after ad

When considering the breadth of particularly economic-based household-level data, there is a vast range of sources. Just in Pakistan, the Labour Force Surveys (LFS), Pakistan Social and Living Standards MeasurementSurveys, Demographic and Health Surveys, to name a few, are run on a regular basis. Not only do these record characteristics such as age, gender, occupation and education level of all household members, most also contain extensive details on consumption and asset ownership. Moreover, over time, these and other surveys have been expanded to include information about such household aspects as health status and outcomes, decision-making, time-use, etc. Such surveys then provide invaluable information on the daily lives of households across Pakistan. However, while there is much that researchers can learn about the opportunities available to and the challenges faced by the population, there remain issues in these data sources as well.

A prime example is that despite being the main source of information on the Pakistan labour market, the LFS only partially captures the scope of informal labour and is especially troublesome when looking to determine the full extent of women’s work. Part of the issue lies in the nature of the questions asked, while the rest lies in inherent enumerator biases. Work by colleagues has shown that those fielding the LFS questionnaire tend to note women as non-working if they are at home during work hours. This is problematic given the tendency of women in Pakistan to work from home even when employed in the formal sector.

In the same vein, standard questions used in surveys, especially those that are fielded in several countries simultaneously, are unable to capture nuances that relate to culture-specific contexts. Thus, it is essential to be mindful of the biases those who are asking the questions as well as the questions themselves may be bringing into the data.

Quantitative data, such as the type that we have become accustomed to seeing, then has its limitations. So, what is the way forward? One is of course to recognise and rectify these biases: gender-sensitised enumerator training and context-based revisions of surveys are essential.

On the one hand, surveys that look to gather more objective data are easier to field and interpret, and can typically cover much larger samples of populations. However, they also give an incomplete snapshot with few of the intricacies that more qualitative methods such as ethnographies, entailing far more detailed conversations with and observation of individuals and communities, provide. The latter, by necessity can only be done for very limited numbers. In an ideal scenario, we would be able to combine the two. Yet, this comes down to the proclivities of the researcher, and inevitably the available funding.

*The writer is assistant professor, economics and director, Saida Waheed Gender Initiative, Lums.*

*Published in Dawn, October 5th, 2020*