

A Tale of Two Methodologies



A Srija & Jitender Singh

For an evidence-based policy, high-frequency, timely, accurate and granular data are critical. The two main sources of employment data for the Economic Survey are the Periodic Labour Force Survey (PLFS) by the National Statistical Office (NSO) and the Consumer Pyramids Household Survey (CPHS) by the Centre for Monitoring Indian Economy (CMIE). While the PLFS quarterly survey, launched in April 2017, covers the urban sector, the CPHS data's frequency is both monthly and weekly, and is available for both rural and urban areas.

As the survey methodologies are different for PLFS and CMIE, it is useful to compare the datasets and undertake some simple smell tests. For the periods for which both datasets are available, in Figure 1, we find a strong negative correlation of -0.75 between PLFS and CMIE data.

Given this negative correlation, how do we decide which dataset captures the real picture of the labour market? One way is to undertake a simple smell test. Okun's law, which relates unemployment rate with the state of the economy, states that employment rises when the economy is doing well and falls when the economy is not doing well. Thus, there should be a positive correlation between the state of the economy and employment rate.

To capture the state of the economy, we use the quarterly GDP at current prices, and correlate it with PLFS and CMIE data. Figure 2 shows that the PLFS measure of employment rate correlates strongly and positively with economic activities (0.84).

In contrast, we see in Figure 3 that the employment rate captured by CMIE correlates strongly, but in a negative way, to economic activities. This does not make economic sense and casts serious doubts on whether the survey methodology employed by CMIE incorporates the necessary scientific rigour without which it cannot be used for research or policymaking.

Strikingly, CPHS estimates the unemployment rate based on just one question — status of employment as on the date of the survey. However, PLFS estimates employment status using a detailed structured questionnaire. Most importantly, there are differences in their

definition and sampling methodology. In CPHS, a person is categorised as employed if he is engaged in any economic activity on the day of the survey or on the preceding day (in case status is uncertain, on the day of the survey).

The employed also include persons who report employed but 'not at work' on the day of the survey, such as a person on leave, on an off-day, a businessman with operative business, person on probation or training. This one-day reference period, therefore, does not care about what a person had been doing in the preceding 4-5 days of the week.

In PLFS-CWS (current weekly status), a person is counted as employed, even if he or she is not employed on the day, or preceding day, of the survey, provided he or she worked at least one hour on any day during the preceding seven days of the survey date. As sampling errors decrease with the increase in the reference period, this key difference

in the reference period seems to be the source of the significant errors in the CMIE data. The scientific rigour of the sampling methodology — including the reference period and the nature of the questionnaire — needs to be significantly refined for the CMIE data to be accurate enough for research and policy.

While PLFS data is very reliable, the significant lag in producing it renders it ineffective for real-time policy changes geared towards employment. For instance, in February 2021, estimates only up to the January-March 2020 quarter are available. Thus, given the significant questions about the CMIE data that render any inference highly questionable, India needs the PLFS infrastructure to be ramped up for regular *monthly* household surveys to track labour market characteristics.

Srija is economic adviser, and Singh is joint director, Department of Economic Affairs, GoI

