Choosing Villages: Sampling Strategy

Dr. Wilima Wadhwa

The purpose of the ASER 2006 rapid assessment survey in rural areas is twofold: (i) to get reliable estimates of the status of children’s schooling and basic learning (reading, writing and math ability) at the district level; and (ii) to measure the change in these basic learning and school statistics from last year. In addition to the basic learning tools that we introduced last year, we have added some higher level tools to the questionnaire this year that tests the comprehension skills of the children. It is also well known that the mother’s education level has an important effect on the child’s educational status as well as learning levels. Therefore, in ASER 2006 we will also have questions on mother’s educational levels and also actually test the mother for basic reading.

The sampling strategy used helps to generate a representative picture of each district. The aim is to survey all rural districts. The estimates obtained will then be aggregated (using appropriate weights) to the state and all-India levels.

Last year, using standard sampling techniques, a sample size of 400 households was derived for each district. This year the sample size is enlarged to 600 households to get an even more accurate account of how the children in each class are faring. Hence, in ASER 2006 Rural, the sample size is 30 villages per district and 20 households per village.

The villages were randomly selected using the village directory of the 2001 census. The sampling was done using the PPS (Probability Proportional to Size Sampling) technique. The PPS is a widely used standard sampling technique and is the appropriate technique to use when the sampling units are of different sizes. In our case, the sampling units are the villages. This method allows villages with larger populations to have a higher chance of being selected in the sample.

In ASER 2006, we kept the 20 villages we had in ASER 2005 and added 10 more villages from the census village directory. The 10 new villages are also chosen using PPS. The 20 old villages and the 10 new villages will give us a “panel” of villages, which generates more precise estimates of changes. Since one of the objectives of ASER 2006 is to measure the change in learning since 2005, creating a panel is a more appropriate sampling strategy. Each district received a village list with appropriate block information along with the data from the 2001 census on total number of households and total population.

Like ASER 2005, the village list is final and cannot be replaced. This is to maintain randomness of the sample to obtain reliable estimates.

The ASER 2006 team has consulted with national level sampling experts including those at NSSO and ISI.