

## **Response Rates and Nonresponse Bias Analysis**

### **School Response Rate**

We plan to select a sample of 3,150 students from a sample of 90 schools in 9 Census divisions. Because this is a longitudinal survey, we anticipate the response rate to be lower than a one-time survey. Therefore, we selected a sample of 135 schools. This includes 90 schools that we plan to contact and a reserve sample of 65 schools. This avoids the problem of going back to the sampling frame for the selection of schools in each Census division in case of schools not agreeing to participate in the survey.

We will make every effort to get the 90 schools in the first sample to participate in the survey. We do not expect to use all the 135 schools selected for the survey. We expect to get a higher than 67% school response rate mentioned in response to the question on school response rate. We hope to achieve a response rate of 80% for the school survey.

For the sampling of students at the second stage, we plan to select an entire class or classes in selected schools. We planned the sample size of students that we wanted to contact assuming a response rate of 80%. This is to have enough students in the sample in case there is a higher nonresponse than one-time surveys because of longitudinal nature of the survey. The assumed response rate is lower than what has been achieved in previous surveys in schools that agreed to participate. Since the entire class is selected and not a sample of students in the class, it is likely, that the response rate for the students for agreeing to participate will be more like 95% than 80%. Response rate for those who agree to participate is expected to be even higher than the assumed initial numbers, As indicated earlier, the assumed response rates are more for estimating the required initial sample size. We expect the overall student response rate to be around 72% and not the very low response rate used to determine initial sampling.

### **Nonresponse Bias Analysis.**

We plan to do a nonresponse bias analysis as per the NCES guidelines.

### **School Nonresponse Bias**

Two components of the bias in the estimates because of nonresponse are the nonresponse rate and the difference between the responding and nonresponding schools with regard to the students' characteristics that are of interest in the survey. It is reasonable to assume that the reasons for school nonresponse to this survey are not related to the characteristics of interest. In other words, the difference between the responding and nonresponding schools with the regard to student health characteristics is not expected to be large. The

nonparticipation by schools is more for administrative and other reasons. We will record the reasons for nonresponse by schools not agreeing to participate in the survey.

We will look at the distribution of nonresponding schools by rural and urban and also by Census divisions or regions to see whether nonresponse rates are higher for rural or urban schools and also in any one Census division or region. We will also look at response rates in public, private and Catholic schools.

In addition to response rates, we will also look at available frame characteristics of schools like enrollment in grade 10 in urban and rural schools in the sample across and within Census divisions. We will also look school characteristics for those schools that readily agree to participate and those schools that first refused and later agreed to participate in the survey.

If the response rates are different for rural and urban schools, we will examine the differences in some important students' characteristics in responding urban and rural schools to see whether this difference in nonresponse rates is introducing any bias in the survey estimates. This analysis may help in determining appropriate nonresponse adjustments to school sampling weights.

#### Student Nonresponse

For the analysis of possible reasons for nonresponse to the survey by students, we plan to record the reasons for nonresponse. This may be due to the longitudinal nature of the survey or related to characteristics of interest in the survey. We plan to keep track of these rates by various subgroups like students in rural schools versus urban schools.

The SSL in each school will provide student characteristics within each classroom which will enable us to examine potential differences between responding and nonresponding students. We will follow-up nonresponding students to see whether these students are different with respect to important characteristics of interest. It will be done on a subsample of students if not the entire sample of nonrespondents.

We also examine whether it is possible to do propensity score analysis if there is sufficient information on predictor variables for both for responding and nonresponding students. As mentioned elsewhere, Dr. Paul Albert, Branch Chief of the Biostatistics and Bioinformatics branch, is an expert in the area of missing data and subject nonresponse and will provide expert assistance in determining the appropriate analyses.

We plan to adjust the sampling weights of responding schools and students to account for nonresponding schools and students. We will look at the possibility of forming weighting classes to adjust the weights.