NAEP National (Math & Science 1996 & 2000 - Grades 4 & 8)
NAEP National (Math 1996 & 2000, Grade 12)
NAEP National (pre-2013)
NAEP National (Reading 1998 & 2000 - Grade 4)
NAEP National (Reading 1998, Grades 8 & 12)
NAEP National (Vocabulary 2009)
NAEP State (Math & Science 2000)
NAEP State (pre-2013)
NAEP State (Reading 1998)

Appendix G: Calculating Design Effects

Surveys in education rarely sample students by simply selecting a random sample of students (known as a simple random sample) from the population. In its simplest form, a sampling design is used where schools (clusters) are first selected and, within each selected school, classes (clusters) or students are randomly selected. Sometimes, geographic areas are first selected before sampling schools and students. The resulting samples are often the result of multi-stage and stratified selection as well.

Students selected from the same school cannot be considered as independent observations as assumed with a simple random sample because they are usually more similar to one another than to students attending other schools.

A simple random sample of 4 000 students is thus likely to cover the diversity of the population better than a sample of 100 schools with 40 students selected from within each school. It follows that the uncertainty associated with any population parameter estimate will be larger for a clustered sample estimate than for a simple random sample estimate of the same size.

In the case of a simple random sample, the variance of a mean estimate $\hat{\mu}$ is calculated as:

$$Var_{SRS(\hat{\mu})} = \frac{\sigma^2}{n}$$

where σ^2 denotes the variance of the whole student population and n is the student sample size. The variance for the mean from a simple random sample is inversely proportional to the number of selected students.

It is usual to express the effect of the sampling design on the standard errors by a statistic called the „design effect". The design effect is calculated as the ratio of the variance of the statistic obtained assuming a complex sample design, to the variance of the same estimate assuming a simple random sample selection of the same number of sampling units (Var_{SRS}). The estimate of the variance for the complex sample is typically calculated using a replication procedure like jackknife or balanced repeated replication ($Var_{REP}(t)$).

Therefore, a design effect for a statistic t can be computed:

$$Deff(t) = \frac{Var_{REP}(t)}{Var_{SRS}(t)}$$