

Sampling Errors

The sample selected for each SPD panel is a stratified multistage probability sample. This complex sample design needs to be taken into account when estimating the variances of the SPD estimates. The SPD data files contain variables, related to the sample design, that are created for the purpose of variance estimation. Several software packages are now available for computing variance estimates for a wide range of statistics based on complex sample designs. Using the variables that specify the design, these programs can calculate appropriate variances of survey estimates. The Census Bureau also provides generalized variance functions (GVFs) that can be used to obtain approximate estimates of sampling variance for the SPD estimates. Information on these functions may generally be found in the technical documentation associated with the data files.

A common mistake in the estimation of sampling errors for survey estimates is to ignore the complex survey design and treat the sample as a simple random sample of the population. This mistake occurs because most standard software packages for data analyses assume simple random sampling for variance estimation. When applied to the SPD estimates, SRS formulas for variances typically underestimate the true variances.