Estimation Process

Each year, AHA’s Health Forum administers the AHA’s Annual Survey. The purpose of the survey is to collect utilization, financial and personnel information on each of the nation’s hospitals. The survey’s overall response rate exceeds 75% each year, which is considerably high for a voluntary survey given its length and the size of the universe (about 6,400 hospitals).

For hospitals that do not respond at all, or not respond fully to the survey, a statistical methodology is run against their records to estimate missing values. The following describes the two major approaches used for estimation:

For nine key variables - total admissions; total births; total inpatient days; total expenses; total full-time employees; total surgical operations; total outpatient visits; total part-time employees; and total revenue – estimates are generated from regression models. For each variable, the previous year’s data (base year) along with estimation status, percentage change in state median, MSA size and bed size are used as the independent variables, while the current year’s value is the dependent variable. The regression model generates a coefficient for each independent variable, which later will be used in estimating the current year’s value. In other words, the current year’s missing value is “predicted” by multiplying the base year data with the corresponding coefficients derived from the regression model.

Alternatively, components of the key variables and all other variables eligible for estimation are generated from a matrix of estimators. An estimator is a ratio of two variables: numerators are the variable to estimate; denominators are an indicator variable such as beds, bassinets or a total variable in which the numerator is an additive component. In the first step, the universe of hospitals is stratified into 52 strata or cells. The stratification is based on the intersection of the following variables, control, service, bed size, and length of stay. Estimators are calculated using all reported data in each stratum. Each estimator takes on a different value for each stratum (thus the matrix). Using payroll expense as an example, matrix estimation occurs as follows. For hospitals in the same stratum, payroll expense as a percentage of total expense is computed based on reported data only. For a non-responding hospital in this same stratum, the estimated payroll expense value is derived by applying the above mentioned percentage to the non-respondent’s total expense value regardless of whether that value is reported or estimated.

Please note that missing facilities and services collected through Section C of the survey are never estimated when missing. Beds, control, service, length of stay also are not estimated but are present in all cases. Values for these variables come from the AHA master facility inventory system. For an institution to be listed in the inventory file there must be a value for those variables. The compilation and management of the inventory is independent of the Annual Survey process.