Adjustments to Data

Edits outliers:

Subject matter analysts review the reported data, and this review commonly results in the reported data receiving one or more corrections, called edits. These edits are based on enforcing a number of consistency checks, including: internal consistency of the reported data, consistency of the reported data to prior period data of that reporting unit (if that exists), consistency of the reported data from other similar units in either the current or prior cycles of the data, and consistency of the reported data to data from other sources. Simple edits are made by reconciling internal inconsistencies in the reported data. Examples of internal inconsistencies include details of investment not equating to the total amount of investment, and reporting investment data in the wrong units. Inconsistencies in the reported data can also occur due to misunderstandings in the definitions or questions, or misidentifying the proper classification of a type of investment into one of the requested categories.

ACES is a company level survey that selects companies as sampling units and uses the entire company as a tabulation unit. The company level sampling units from the ACE-1 frame, companies having paid employees that were chosen during the probability sampling portion have a sampling weight that is a rough guide to how many companies' data is represented by the selected company. The sampling weight is an algebraic multiplier that allows the response data from one tabulation unit to be used for statistical inference to similar units that were not selected in the sample. In some cases, the response data is correct for that company, but not considered reliably representative of other similar units that were not selected in the sample. In this event, these tabulation units may have their sampling weight marginally reduced, and are identified as outliers.

Outlier detection is done initially through calculating the Mahalanobis distance between investment in structures and equipment. While ACES is primarily concerned with total capital expenditures, the lack of a reliable covariate makes outlier detection on total capital expenditures difficult. The focus is then on the relative values of structures and equipment investment, which is examined by ACES specific industry. The reporting companies with the largest Mahalanobis distance for each ACES specific industry are identified. These are then reviewed as potential outliers by the survey analysts to double check the validity or representativeness of the reported data. Based on the survey analysts' review and recommendation, select companies are inspected against similar companies as well as their impact on the estimates of structures or equipment expenditures in that ACES industry. A company passing those tests is considered an outlier. Its nonresponse-adjusted weight will be reduced in a manner to reduce the sampling variation.

Outliers still represent the entire amount they did before this adjustment, but their impact on the estimate is reduced. This is done in an attempt to improve the representativeness of the sample and improve the quality of this portion of the estimate, when there is evidence of extraordinary investment.

The company level tabulation units in the ACE-2 frame, companies without paid employees, are similar, but there is no general inspection for outliers. However, company level tabulation units from the 3C and 3F component are inspected. These companies are believed to have no associated employees and were removed from the ACE-1 frame. They were removed for either being a possible new company that may or may not have employees (the 3F strata), or a company that formerly had employees but is now believed to not have employees (the 3C strata).

Nonresponse adjustment and imputation:

Capital expenditures have low correlation with administrative or collected data, but show some relation to business activity and general size of the company. Attempting to model capital expenditures based on the administrative or collected data risks an unacceptable level of estimation errors.

If a sampled company does not report, which is known as unit nonresponse, the assumption is that the company has capital expenditures in direct proportion to its payroll measured against other companies that did report in the same ACES specific industry and payroll group. ACES uses a weight adjustment to reporting companies to account for companies that do not report. The weight adjustment for nonresponse is made to the original sampling weight, so that the sample can be used to make inference back to the entire population.

If a company reports capital expenditures sufficiently to be a respondent to the survey, but does not answer a particular question, known as item nonresponse, then either a survey analyst will estimate the missing data or no correction will be made.

The nonresponse weighting adjustment differs between the two samples. Every company in the ACE-1 sample, which are all determined to have employees, has payroll. In addition, every company in the ACE-1 sample has been put into a group of businesses with similar activities. If a company fails to respond to the survey, and that company is still considered to have been active and in scope during the survey cycle, then its payroll is added to other companies in that ACES industry group and substrata that also did not report. Each ACES industry has at least one to as many as five substrata. Each substratum should have sufficient number sampled to ensure there is at least one respondent in the substrata, or separate action has to be taken. The fraction of the sum of the payroll of all sampled companies over the sum of payroll of just responding companies is calculated for each substratum. This value, which must be greater than or equal to one, is multiplied by the sampling weight of each responding company to compute a nonresponse adjusted weight.

Companies in the ACE-2 sample are assumed not to have employees or payroll. Further, the business activity codes associated with them are not considered reliable enough to use for estimation. The compensating fraction for a nonresponse adjustment will be the number of sampled but still eligible cases over the number of eligible reporting cases. This value, which must be greater than or equal to one, is multiplied to the sampling weight of each responding company to compute a nonresponse adjusted weight.

The nonresponse-adjusted weights for both the ACE-1 and ACE-2 samples are used in estimation. Use of the nonresponse weight adjustments assumes companies that do not report are not different from those that do with respect to capital expenditures. The adjusted weights also provide enough influence to responding companies for adequate inference to the population. The higher weight values also increase estimation variability, which should occur with larger non-observation due not only to sampling, but also to decreased response rates.

This survey does not use imputation other than that provided by the survey analysts. Missing data is handled with the nonresponse weight adjustment as detailed above.