Cut-off samples, based on a model-based sampling approach, can be useful for highly skewed establishment surveys with a frequently repeated sample where the target population is also monitored periodically by a census survey or reliable third-party data. Since the respondent data from the sample and census surveys that collect the same data items from the same companies are expected to have high positive correlation, a model-based approach that capitalizes on this relationship typically yields small and efficient samples. This saves cost and reduces respondent burden. The theoretical underpinnings of this methodology were presented in seminal papers by Brewer (1963) and Royall (1970). Essentially, the samples are comprised of relatively large units.

A weighted least squares regression model is typically the underlying superpopulation model used for many establishment surveys at the U.S. Energy Information Administration (EIA). The regression model also forms the basis for the estimation methodology. In addition, this model can be used to predict for nonresponse (and edit failure) values. The advantage of this approach is that model-based variances for survey estimates also account for imputation, and they can be calculated seamlessly using standard regression formulae.

An application of this prediction-based imputation approach is presented for a monthly survey of natural gas suppliers. This survey collects data used for estimating state level natural gas total sales volume and average price published in EIA's Natural Gas Monthly.