Weighting and Variance Estimation Under Responsive Designs and Related Forms of Sample-Driven Resource Allocation for Complex Surveys

John L. Eltinge (Bureau of Labor Statistics)

With some notable exceptions, formal inference from complex sample data has traditionally been based on the premise that randomization mechanisms for resource allocation are determined a priori. This resource allocation may include several components, including sample selection; assignment of collection mode and interviewers; nonresponse follow-up; and re-interviews to evaluate measurement error properties. In practice, however, field-level decisions in the abovementioned areas often involve decisions that are not entirely determined a priori, and that may be based on paradata or other preliminary information available on some of the sample units. For the case of nonresponse follow-up, Groves and Heeringa (2006) and subsequent publications explored some of these decision processes under a "responsive design" framework, and suggested a number of ways in which to use this framework to improve the balance between data quality and cost.

The current paper explores issues of weighting and variance estimation under a "responsive design" framework and related forms of sample-driven resource allocation. It places primary emphasis on methods based on extensions of standard analyses of data collected through two-phase or multi-phase sample designs. Three complementary approaches receive primary attention. These involve methods based on, respectively: (1) a set of randomized resource-allocation rules determined a priori; (2) conditioning on the observed paradata; or (3) integration with respect to the distributions induced by a superpopulation model for the paradata, as well as the randomization distribution induced by the original sample design. A simulation study illustrates some practical distinctions in results obtained through approaches (1) through (3), respectively.