



## **Design, data collection and estimation**

Peter Lundquist\* Ph.D.

Statistics Sweden, Stockholm, Sweden, peter.lundquist@scb.se

Carl-Erik Särndal, Professor Emeritus

Statistics Sweden, Ystad, Sweden, carl.sarndal@telia.com

### **Abstract**

Three important areas in the production of an interviewer administered household survey are the design, the data collection, and the estimation. Survey design and estimation are by tradition assignments for the survey statistician whereas the data collection is left to other methodologists who not always work in the same organization. The survey climate today with increased nonresponse has introduced new strategies such as responsive design (or adaptive design). By these methods it has become more clear that the data collection is an important area for the survey statistician to consider. Interventions are made during data collection with the ambition to produce a well-balanced or well representative set of respondents. Irrespectively if the interventions are based on a plan in advance or based on observations during the data collection, accurate estimation remains the ultimate goal. An objective in this presentation is to examine the question: What benefits can be expected at the estimation stage from having improved the balance of the response set during data collection? Somewhat improved balance does not by itself guarantee estimates with low bias. Is balance worth a perhaps costly and demanding effort at the data collection stage? Or could one have done equally well by saving the use of the auxiliary information until the estimation stage? Complete elimination of the bias is not achieved in the estimation phase either, but the goal is best accuracy (least bias) given the circumstances. Data from a Swedish survey will be used in simulations to illustrate the questions above. The evaluation will be based on recently developed indicators for nonresponse.

**Keywords:** Auxiliary variables, balanced response, nonresponse, responsive design