"Challenging Research Issues in Statistics and Survey Methodology at the BLS"

Topic Statement: Modeling, Estimation and Inference Issues that Arise from the Use of Response-Rate Goals in Survey Operations

Key Words: Incomplete data; Interviewer effect; Missing data; Nonresponse; Pattern-mixture model; Quasi-randomization model; Quota sampling; Survey costs; Total survey error model; Variance estimation.

Contact for further discussion:
John L. Eltinge
Office of Survey Methods Research, PSB 1950
Bureau of Labor Statistics
2 Massachusetts Avenue NE
Washington, DC 20212
Telephone: (202) 691-7404
Fax: (202) 691-7426
E-mail: Eltinge.John@bls.gov

Background, Definitions and Notation:

In practical work with survey data, we often encounter *nonresponse*, in which a selected sample unit does not provide responses to one or more items on the survey data collection instrument. There is a large literature on this topic; see, e.g., Little and Rubin (2002), Groves et al. (2002), Groves and Couper (1998), Madow, Nisselson and Olkin (1983), Madow, Olkin and Rubin (1983), Madow and Olkin (1983), and references cited therein.

Much of this literature is based (implicitly or explicitly) on simple quasirandomization models, in which one defines a response indicator

$$r_i = \{1 \text{ if sample unit } i \text{ responds; } 0 \text{ otherwise}\}, \qquad (Q.1)$$

assumes that the r_i are Bernoulli (p_i) random variables, and then models the response probabilities p_i through logistic regression or other methods involving a fixed set of predictors X_i, say. In addition, much of this literature is based on the assumption that the response indicators are independent across sample units *i*. For some general background on quasirandomization models, see Oh and Scheuren (1983).

However, it appears that due to contractual or regulatory factors, some large survey organizations have goals or incentives tied to achievement of specified response-rate goals, and have relatively little additional incentive to achieve response rates above the specified goals. The effects of such goals can be especially important in panel surveys or other surveys in which there is a relatively brief time available for data collection.

In some cases, the abovementioned goals or incentives are operational only at an overall organizational level, while in other cases similar goals or incentives apply to individual field supervisory staff (e.g., a regional or area office manager), or to individual interviewers. At an extreme, these goals or incentives can lead to forms of quota sampling, with its attendant problems with bias and lack of information for appropriate adjustment of point estimators and variance estimators.

For surveys subject to response-rate goals or incentives, implementation of those goals or incentives may lead to models for nonresponse that are distinct from standard quasirandomization models (Q.1). For example, in some cases it may no longer be plausible to treat response probabilities p_i as dependent only on the fixed predictors X_i; and it may no longer be plausible to treat the response indicators r_i as independent across sample units. To illustrate with an over-simplified example, suppose that for a given target population the following conditions hold.

(i)   Within a specified stratum, data collection efforts begin with an attempt to contact and interview each sample unit i ; on this initial attempt, response indicators r_i are independent and follow model (Q.1).

(ii)   Define q_i to be the probability that a single nonresponse follow-up attempt will result in a completed interview, conditional on the interview staff attempting to re-contact the nonresponding sample unit i. In addition, assume that the terms q_i are monotone increasing in the initial response probabilities p_i. Then under mild conditions, one would maximize the overall response rate for a given investment of follow-up resources by concentrating nonresponse follow-up efforts on nonrespondents that had relatively high probabilities p_i of initial response.

Concentrating nonresponse follow-up efforts in the form suggested in (ii) would lead to several issues in point estimation and inference from the resulting survey data, including the following.

(A) The overall unconditional probability of response for a given unit i would depend on: the probabilities p_j and q_j for each unit j in the sample; the response-rate goal established for the group containing unit i; and the specific targeting strategy used in nonresponse follow-up work.

(B) Standard methods of nonresponse adjustment (e.g., construction of weighting or imputation cells based on simple classificatory variables, or based on response probabilities estimated from an unconditional logistic regression model) may not fully account for the conditional-probability structure induced by the follow-up methods in (ii).

2

In practice, nonresponse follow-up efforts are more complex than suggested by (i)-(ii). For example the true initial-response probabilities p_i are not known. In addition, incentives can be structured to encourage efforts to collect information from units with relatively low initial response probabilities p_i, e.g., by establishing response-rate goals separately within groups that have, respectively, high and low initial response probabilities p_i. Also, there often will be more than one follow-up attempt for a given nonresponding sample unit.

Issue: What are appropriate ways in which to account for the impact of response-rate goals or incentives in development of nonresponse-adjusted point estimators and inference methods?

Questions on Nonresponse Adjustment Methods and Related Methodological Work in the Presence of Response-Rate Goals or Incentives:

1. First consider the case in which response goals or incentives are administered only at an institutional level.

    (a) What are appropriate ways in which to account for response-rate goals or incentives in the construction of nonresponse-adjusted point estimators, e.g., estimators based on weighting adjustment or imputation?

    (b) What are appropriate ways in which to account for response-rate goals, and the adjustment methods in (a), in construction of variance estimators and inference methods?

    (c) In some survey fieldwork, response-rate goals or incentives depend in part on the types of nonresponse encountered by an interviewer. For example, noncontacts or refusals may "count against" the interviewer, while sample units that no longer exist or are out of scope do not "count against" the interviewer. For these cases, do we need to modify or expand our answers to (1.a) and (1.b)?

2. Now consider the case in which response goals or incentives are administered at the level of an individual interviewer. For example, each interviewer may be expected to complete a specified percentage of assigned sample cases. Then, in addition to the issues identified in question (1), the interviewer-level goals or incentives may induce an "interviewer effect" in the nonsampling error component of a total survey error model.

    (a) What are appropriate models through which to incorporate this specific type of interviewer effect into a total survey error model?

(b) Under the conditions described through the models in (a), what are appropriate point estimators and variance estimators that account for both sampling and nonsampling error components, including the abovementioned interviewer effect? For this question, a simplifying assumption would be that interviewers are assigned randomly to sample units.

(c) As an extension of the simplified case considered in (b), anecdotal evidence indicates that the assignment of interviewers to sample units is not entirely random in some cases. Instead, sample units that are difficult to contact or "reluctant" (e.g., sample units that have not responded to initial interview attempts) may be assigned to interviewers who have especially high levels of training or experience. In an informal sense, these special interviewers are believed to have a higher probability of "converting" the sample unit to respondent status. What are appropriate models for the (partially) nonrandom assignment of interviewers to sample units in the presence of response-rate goals or incentives? What adjusted point estimation or inference methods follow from those models?

(d) The abovementioned issues arise in a relatively simple form when a given set of interview cases are assigned to a single interviewer, who has sole responsibility for the final disposition of these cases through a personal visit or telephone interview. More complex forms of these issues may arise if nonresponding units are referred to refusal conversion specialists.

(e) If response-rate goals and incentives are applied separately for individual interviewers and for supervisory staff, one would anticipate the use of a hierarchical form for the resulting total survey error model. What are the specific ways in which the multiple levels of response-rate goals should be incorporated into the hierarchical model?

(f) In the literature on point estimation and inference from survey data subject to nonresponse, large-sample approximations generally are based on conditions that do not account explicitly for the response-rate goals described here. To what extent, if any, does one need to change standard asymptotic conditions to reflect the use of response-rate goals? Do the changed conditions lead to substantial changes in the development of asymptotic results for point estimation and inference methods in the presence of response-rate goals?

3. In analyses of incomplete data, some authors consider point estimation and inference methods that provide different treatment of data from "early reporters" and "late reporters," respectively, where there may be callbacks or other follow-up efforts with nonrespondents after an initial response period. See, e.g., Drew and Fuller (1981, 1982), Merkle et al. (1993) and Potthoff et al. (1993). These estimation and inference methods generally are developed under selection models

4

similar to those in Question 1, but possibly with response probabilities differing according to the number of callback attempts received to date.

To what extent can this "callback-based" estimation literature be applied or extended to the conditions described in Questions 1 or 2?

4. Questions 1 through 3 focused on selection models for nonresponse. *Pattern-mixture models* provide an alternative approach to nonresponse analyses and adjustments. For some general background, see, e.g., Little (1993, 1994). In general, pattern-mixture models are of special interest for cases in which:

   (a) our data have panel or other multivariate structure; and

   (b) units with the same multivariate pattern of response and nonresponse can reasonably be expected to have a common mean and covariance structure.

Pattern-mixture models often are considered to be of special interest for certain types of nonignorable nonresponse.

To what extent, if any, do the nonresponse models from Questions 1 and 2 lead to pattern-mixture models, and related estimators, that differ substantially from those developed previously in the pattern-mixture literature?

References:

Binder, D.A. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review* **51**, 279-292.

Drew, J.H. and Fuller, W.A. (1980), Modeling nonresponse in surveys with callbacks, *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 639-642

Drew, J.H. and Fuller, W.A. (1981), Nonresponse in complex multiphase surveys, *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 623-628

Groves, R.M. and Couper, M.P. (1998). *Nonresponse in Household Interview Surveys.* New York: Wiley.

Groves, R.M., D. Dillman, J.L. Eltinge and R.J.A. Little, eds. (2002). *Survey Nonresponse.* New York: Wiley.

Kennickell, A.B. (2000). Asymmetric information, interviewer behavior and unit nonresponse. Paper presented at the Joint Statistical Meetings, August, 2000.

Little, R.J.A. (1993). Pattern-mixture models for multivariate incomplete data, *Journal of the American Statistical Association*, **88**, 125-134

Little, R.J.A. (1994). A class of pattern-mixture models for normal incomplete data', *Biometrika*, 81 , 471-483.

Little, R.J.A. and D.B. Rubin (2002). *Statistical analysis with missing data*, New York: Wiley.

Madow, W. G., Nisselson, J. and Olkin, I., eds. (1983). Incomplete data in sample surveys, volume 1: Report and case studies. New York: Academic Press.

Madow, W. G.;Olkin, I., and Rubin, D. B., eds. (1983). Incomplete data in sample surveys, volume 2: Theory and bibliographies. New York: Academic Press.

Madow, W. G. and Olkin, I., eds. (1983). Incomplete data in sample surveys, volume 3: Proceedings of the symposium. New York: Academic Press.

Merkle, D.M., S.L. Bauman and P.J. Lavrakas (1993). The impact of callbacks on survey estimates in an annual RDD survey, *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 1070-1075

Oh, H.L. and Scheuren, F.J. (1983). Weighting adjustment for unit nonresponse. Pp. 143-184 in W.G. Madow, I. Olkin and D.B. Rubin, eds., *Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliography.* New York: Academic Press

Potthoff, R.F., K.G. Manton and M.A. Woodbury (1993), Correcting for nonavailability bias in surveys by weighting based on number of callbacks, *Journal of the American Statistical Association*, **88**, 1197-1207

Rao, J.N.K. and Shao, J. (1992), Jackknife variance estimation with survey data under hot deck imputation, *Biometrika*, **79**, 811-822

Rosen, B. (1972a), Asymptotic theory for successive sampling with varying probabilities without replacement, I, *The Annals of Mathematical Statistics*, **43**, 373-397

Rosen, B. (1972b), Asymptotic theory for successive sampling with varying probabilities without replacement, II, *The Annals of Mathematical Statistics*, **43**, 748-776