OPTIMAL SURVEY DESIGN WHEN NONRESPONDENTS ARE SUBSAMPLED FOR FOLLOWUP

A. James O'Malley and Alan M. Zaslavsky
Department of Health Care Policy, Harvard Medical School,
180 Longwood Avenue, Boston, MA 02115-5899 (omalley@hcp.med.harvard.edu)

Key Words: Neyman allocation, Nonrespondents, Small Area Estimation, Subsample, Survey, Telephone Followup.

Abstract:

Surveys often first mail questionnaires to sampled subjects and then follow up mail nonrespondents by phone. The high unit costs of telephone interviews make it cost-effective to subsample the followup. We derive optimal subsampling rates for the phone subsample for comparisons of health plans or other units. Computations under design-based inference depart from the traditional formulae for Neyman allocation because the phone sample size at each plan is constrained by the number of mail non-respondents and multiple plans are subject to a single cost constraint. Because plan means for mail respondents are highly correlated with those for phone respondents, more precise estimates (at fixed overall cost) for potential phone respondents are obtained by combining the direct estimates from phone followup with predictions from the mail survey using small-area estimation (SAE) models.

1 Introduction

In Consumer Assessments of Healthcare Providers and Systems (CAHPS®) surveys, randomly selected members of health plans are sent questionnaires by mail, along with a prepaid return envelope (Crofton, Lubalin, and Darby 1999, Goldstein, Cleary, Langwell, Zaslavsky, and Heller 2001). Those that do not return their survey are sent a reminder notice and if they still do not respond four attempts are made to reach them by phone. From the survey methods standpoint, there are at least three distinct populations: mail responders, phone-only responders, and nonresponders. Ideally we want to obtain results for the entire population but unless the nonresponders are interviewed by some other means no information is available for them. The population of those who would respond either by mail or phone is thus our population of focus. The estimates of the plan means for this population are the basis of comparisons of quality between plans.

In surveys like those in CAHPS, phone respondents will differ from mail respondents whenever individual characteristics affect the likelihood of responding by mail, and therefore estimates are required for both sub-populations. Because the telephone survey is typically more expensive than the mail survey per response obtained, it might be cost-effective to subsample in the follow-up survey. CAHPS mail non-respondents were subsampled for the first time in 2006, but the design was not formerly optimized to minimize costs.

In this paper we consider optimal designs for surveys of this type. We first compute the optimal sample size using design-based inference. The computation departs from the traditional formulae used for Neyman allocation (Cochran 1977) because the phone sample size at each plan is bounded above by the number of mail non-respondents. We consider several objective functions based on the variances of the estimators.

Because the plan means for mail respondents tend to be highly correlated with those for phone respondents, more precise estimates for the phone respondents might be obtained by combining the prediction from the mail survey with the direct estimates from the phone followup survey. We use small-area estimation (SAE) models to estimate means for the population of respondents that do not respond by mail but would respond to the telephone sub-sample. SAE allows improved estimates to be obtained for domains with small sample sizes and thus has the potential to obtain more efficient estimates across the analysis. We compare the variance estimates and the associated optimal designs between the direct and SAE estimation strategies.

Sample size calculations for a SAE model are often more complex than those for the corresponding design-based approach. We present an approximate calculation which assumes the response rates and model parameters are known. This sets the scene for more exact calculations in future work.

In general, the methods developed in this paper can be extended to include a range of covariate effects. However, in our application only survey de-

Supported by grant U18 HS09205-06 from the U.S. Agency for Healthcare Research and contract 500-95-007 from the U.S. Centers for Medicare and Medicaid Services.

scriptors such as response rates are available.

2 Notation and Objectives

Let y_{i1}, θ_{ij}, σ^2_{ij}, r_{ij}, π_{ij}, and n_{ij} denote the sample mean rating, population mean rating, population variance of ratings, sample response rate, probability of response, and sample size for an item using mode j (0 = mail, 1 = phone) at plan i. In this paper we ignore uncertainty in the response rates by assuming $\pi_{ij} = r_{ij}$. The number of respondents is given by $m_{ij} = r_{ij}n_{ij}$.

The overall mean for plan i is given by $\eta_i = \delta_i\theta_{i0} + (1-\delta_i)\theta_{i1}$ where δ_i is the proportion of all potential respondents at plan i that would respond by mail. This can also be expressed as the mail mean plus an adjustment for telephone mode:

$$\eta_i = \theta_{i0} + (1-\delta_i)(\theta_{i1} - \theta_{i0}). \quad (1)$$

Because sampled individuals decide whether to respond to the mail survey, to subsequently respond by phone, or not to respond at all, $\theta_{i1} - \theta_{i0}$ is not the "mode effect" in the usual survey methodology sense but rather is a combination of effects of mode and population selection.

Let d_{ijk} denote the cost of sampling an individual at plan i using mode j when their response status is k (1 = responds, 0 = does not respond). Because a phone interview is likely to involve more human time than a non-response or a refusal and returned surveys are likely to involve more processing time than non-returned surveys it is likely that $d_{i0k} < d_{i1k}$ and $d_{ij0} < d_{ij1}$. We present the general case where costs and variances can vary across plans but note that in most applications one would expect the cost and variances to be constant across plans. For the design-based and SAE calculations π_{i0} and π_{i1} are assumed known when in reality they are estimated from the previous years results. The uncertainty about π_{i0} and π_{i1} will be incorporated in future work.

The average costs per individual sampled via mail or phone at plan i are weighted averages of the cost per respondent and cost per non-respondent, given by $c_{i0} = \pi_{i0}d_{i01} + (1-\pi_{i0})d_{i00}$ and $c_{i1} = \pi_{i1}d_{i11} + (1-\pi_{i1})d_{i10}$ respectively. The cost of the total sample at plan i is $n_{i0}c_{i0} + n_{i1}c_{i1}$. The problem considered here is to minimize the total cost of the survey

$$\text{cost}(n) = \sum_i n_{i0}c_{i0} + n_{i1}c_{i1}, \quad (2)$$

subject to the precision constraint:

$$\text{var}(\hat{\eta}_i) = (\delta_i, 1-\delta_i)\text{cov}(\hat{\theta}_i)(\delta_i, 1-\delta_i)^T \le v_0 \ \forall i, \quad (3)$$

where $v_0 > 0$ is a pre-specified reporting standard. Because any slack in the constraints (3) entails increased cost, we may replace the inequalities in (3) with equalities.

A second set of constraints arise because the number of members sampled by phone cannot exceed the number of non-respondents by mail,

$$n_{i1} \le n_{i0}(1-\pi_{i0}) \ \forall i. \quad (4)$$

Equation (4) will be binding (satisfied with equality) if the optimal solution would otherwise require sampling more phone respondents than mail nonrespondents.

3 Sample Size Calculation for the Design-based Estimator

.

We suppose that individuals are sampled using simple random sampling within each plan. Because we are interested in the super-population of all current and potential future plan members we do not make finite population corrections. The estimators of the mail and phone means are $\hat{\theta}_{i0} = y_{i0}$ and $\hat{\theta}_{i1} = y_{i1}$ respectively and the variance of $\hat{\eta}_i$ is

$$\text{var}(\hat{\eta}_i) = \delta_i^2\text{var}(\hat{\theta}_{i0}) + (1-\delta_i)^2\text{var}(\hat{\theta}_{i1}),$$

where $\text{var}(\hat{\theta}_{ij}) = \sigma^2_{ij}/(n_{ij}r_{ij})$. To simplify the following presentation we let $w_{i0} = \delta_i^2\sigma^2_{i0}/r_{i0}$ and $w_{i1} = (1-\delta_i)^2\sigma^2_{i1}/r_{i1}$.

The optimal solution is obtained by finding the optimal phone/mail ratio and then the sample sizes that correspond to this ratio. If (4) is binding $n_{i1} = n_{i0}(1-r_{i0})$ and so the optimal ratio is $1-r_{i0}$. If (4) is slack the method of Lagrange multiplers reveals that $n_{i0} = \lambda_i w_{i0}/c_{i0}$ and $n_{i1} = \lambda_i w_{i1}/c_{i1}$, where λ_i is the Lagrange multiplier for the ith precision constraint, the standard result for optimal allocation in stratified sampling, Therefore, the optimal ratio encapsulating both two scenarios is

$$t_i = \min\left\{1 - r_{i0}, \left(\frac{w_{i1}c_{i0}}{w_{i0}c_{i1}}\right)^{1/2}\right\}.$$

The ratio of mail to phone surveys is t_i^{-1} which implies that the relative number of mail surveys increases with δ_i and σ^2_{i0} but decreases with c_{i0}/c_{i1} and r_{i0}. Substituting $n_{i1} = t_i n_{i0}$ into (3) we obtain

$$n_{i0} = \frac{1}{v_0}\left(w_{0i} + \frac{w_{1i}}{t_i}\right),$$

$$n_{i1} = \frac{t_i}{v_0}\left(w_{0i} + \frac{w_{1i}}{t_i}\right).$$

In the Appendix we derive the same result using a Lagrange multiplier representation of the full optimization problem.

4 Small Area Estimation Model

To improve the precision with which the phone means are estimated we use a small-area estimation model of the mail and the phone means. This enables the information in the phone predictions from the mail survey, which varies depending on the level of correlation between mail and phone, to be combined with the direct estimates of the phone means.

To account for the uncertainty in the measurement of both the mail and phone means, we use the following multivariate model:

$$y_{ij} \mid \theta_{ij} \sim N(\theta_{ij}, \sigma_{ij}^2/m_{ij})$$

where

$$\begin{pmatrix} \theta_{i0} \\ \theta_{i1} \end{pmatrix} \sim N\begin{pmatrix} x_{i0}^T\beta_0 \\ x_{i1}^T\beta_1 \end{pmatrix}, \Sigma \Bigg),$$

x_{ij} is a vector of plan-model level predictors, and σ_{ij}^2 is the residual variance. The correlation coefficient $\rho = \Sigma_{01}/(\Sigma_{00}\Sigma_{11})^{1/2}$ measures the association between the mail and the phone means. The variances Σ_{00} and Σ_{11} measure the amount that the mean ratings deviate from the regression equations.

In our problem $x_{ij}^T = (1, r_{i0})$. As for the design-based calculations we treat r_{i0} as measured without error (i.e., as if $r_{i0} = \pi_{i0}$). In general we could include other domain-level covariates in x_{ij} such as the average number of mail questionaires sent and the average number of callbacks needed to complete a phone survey. However, in the current data the mail response rate is all that is available.

4.1 Posterior Variance

Under Bayesian analysis, η_i is estimated by its posterior mean and the posterior variance is used in the optimal sample size calculations. The posterior distribution of $\theta_i \mid \beta, \Sigma, Y_i, X_i$ is $N(P_i\bar{Y}_i + (I - P_i)X_i^T\beta), (V_i^{-1} + \Sigma^{-1})^{-1})$, where $P_i = (V_i^{-1} + \Sigma^{-1})^{-1}V_i^{-1}$, implying that the posterior variance of η_i is given by

$$\text{var}(\eta_i \mid \beta, Y, X) = (\delta_i, 1-\delta_i)(V_i^{-1} + \Sigma^{-1})^{-1}(\delta_i, 1-\delta_i)^T. \quad (5)$$

A numerical procedure (described later) may be used to find the n_i that minimizes the cost of the survey while satisfying the precision requirement based on (5).

5 Composite estimator

Because the mean ratings for mail respondents are based on larger sample sizes and therefore likely to be estimated more accurately than means for phone respondents, SAE is likely to have less impact on the mail means. Furthermore, the model-based estimate has the potential to introduce bias whereas the direct estimate is design-unbiased, a characteristic that is traditionally considered desirable for CAHPS estimation. Thus, in practice it may be desirable to combine the design-based mail estimate with the model-based estimate of the difference $\theta_{Di} = \theta_{i1} - \theta_{i0}$ between the mail and phone means:

$$\hat{\eta}_i = \bar{y}_{i0} + (1-\delta_i)E[\theta_{Di} \mid \beta, Y, X]. \quad (6)$$

5.1 Variance of Composite Estimator

Because $E[\theta_{Di} \mid \beta, Y, X]$ depends on the data through $\bar{y}_{i1} - \bar{y}_{i0}$, the variance of the $\hat{\eta}_i$ in (6) is computed with respect to the sampling distribution of $\bar{y}_i$. Specifically,

$$\begin{aligned} E[\theta_{Di} \mid \beta, Y, X] &= \quad (7) \\ &(1,-1)(P_i\bar{Y}_i + (I-P_i)X_i^T\beta) \\ &= d_{i0}(n)\bar{y}_{i0} - d_{i1}(n)\bar{y}_{i1} + (I-P_i)X_i^T\beta, \quad (8) \end{aligned}$$

where

$$d_{i0}(n) = \frac{1}{\det(V_i^{-1} + \Sigma^{-1})}\left\{\left(\frac{n_{i1}}{\sigma_{i1}^2} + \Lambda_{11} + \Lambda_{01}\right)\frac{n_{i0}}{\sigma_{i0}^2}\right\},$$

$$d_{i1}(n) = \frac{1}{\det(V_i^{-1} + \Sigma^{-1})}\left\{\left(\frac{n_{i0}}{\sigma_{i0}^2} + \Lambda_{00} + \Lambda_{01}\right)\frac{n_{i1}}{\sigma_{i1}^2}\right\},$$

and $\Lambda = \Sigma^{-1}$.

Therefore, with β and Σ treated as a known constants, the sampling variance of $\hat{\eta}_i$ is given by:

$$\text{var}(\hat{\eta}_i) = (1 + (1-\delta_i)d_{i0}(n))^2\frac{\sigma_{i0}^2}{n_{i0}} + ((1-\delta_i)d_{i1}(n))^2\frac{\sigma_{i1}^2}{n_{i1}}.$$

Because the expression in (8) contains a term that does not depend on Y_i, the bias of $\hat{\eta}_i$ is likely to be non-zero. Therefore, we use the mean squared error (MSE) instead of the sampling variance in the computation of optimal sample size. We have

$$\text{MSE}(\hat{\eta}_i) = \text{var}(\hat{\eta}_i) + \text{bias}(\hat{\eta}_i)^2,$$

where $\text{bias}(\hat{\eta}_i)^2 = b_i^T\Sigma b_i$ and $b_i^T = (1-\delta_i)(1,-1)(I-P_i)$. Thus, $\text{MSE}(\hat{\eta}_i)$ is substituted for the variance term in the evaluation of the precision constraint.

6 Derivation of Model-based Optimal Sample Size

Lagrange multiplier optimization problems may be used to obtain the optimal sample sizes under the full SAE model or the composite estimator. However, because the posterior variance of η_i and $\text{MSE}(\hat{\eta}_i)$ are nonlinear functions of n_i, closed form solutions do not in general exist.

Rather than numerically solving the Lagrange multiplier problem described in the preceding paragraph, we instead first solve the dual problem in which the optimal n_i minimizes the variance subject to a fixed cost for each domain. Because interchanging the objective function with the precision constraint does not change the Lagrange multiplier conditions, at the optimal solution the ratio $t_i = n_{i1}/n_{i0}$ is the same in both problems. Therefore, upon solving the dual problem the only thing left to do is re-scale the sample sizes to satisfy the precision constraint at each domain.

We find it easier to obtain the optimal solution using this two-step process rather than directly solving the original problem because the cost constraint (in the dual problem) is linear in n_i whereas the precision constraint (in the original problem) is nonlinear in n_i. Nonlinear objective functions are less burdensome because the optimal solution is found by searching the line in $\mathcal{R}^2$ for which the constraints are satisfied (brute force or univariate optimization methods can be used).

The computation of the optimal sample size for domain i is summarized by the following algorithm:

1. Generate a right-hand-side for the "cost constraint" by evaluating $k_0 = N^{-1} \sum_i (c_{i0}n_{i0} + c_{i1}n_{i1})$ using as initial values $n_{i0} = n_{i1} = (\sigma_{i0}^2 + \sigma_{i1}^2)/v_0$, where N is the number of plans. Clearly, this is constant across domains and so only needs to be computed once.

2. Determine the feasible region for the (ith) dual problem:

$$\mathcal{F}_i = \{n_i : n_{i1} = \min[(k_i - c_{i0}n_{i0})/c_{i1}, (1 - r_{i0})n_{i0}]\}.$$

3. Search along $\mathcal{F}_i$ to find n_i^{dopt}, the optimal solution to the dual problem.

4. Re-scale the optimal solution to the dual problem to approximate the optimal solution to the original problem:

5. Let $n_i = n_i^{\text{dopt}}$. Repeat the following until convergence:

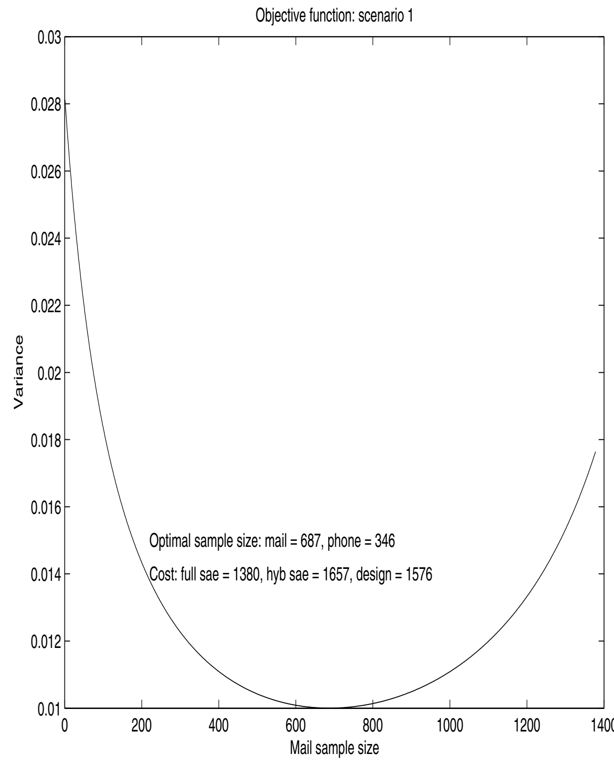

Figure 1: Variance of an estimated plan mean as a function of n_{i0} when $v_0 = 0.01$.

   (a) Compute $q_i = \text{var}(n_i)/v_0$, where $\text{var}(n_i)$ denotes the variance (or MSE) used in the optimization problem in step 3.

   (b) Set $n_i \leftarrow q_i n_i$.

Step 5a is required because the sum of the variances for the mail and the phone estimators is not a linear function of n_i. Thus, re-scaling the solution to the dual problem by the amount that the variance is exceeded (or not) gets very close but does not exactly solve the original problem. However, the number of iterations until the variance constraint is satisfied (to the numerical accuracy of the computer) is small and so convergence is almost immediate.

Step 5b is equivalent to finding the value of k_i at which the objective function of the dual problem equals v_0. The variance when k_i is such that $v_0 = 0.01$ is shown as a function of n_{i0} in Figure 1.

7 Results

To illustrate the extent that the SAE can reduce the required sample size we generated data that resemble data typically seen in CAHPS. We set $v_0 = 0.01$, $c_{i0} = 1$, $c_{i1} = 2$, $\sigma_{i0}^2 = 3.78 = \sigma_{i1}^2 = 3.78$, $\delta = 0.5$, and

$$\Sigma = \begin{pmatrix} 0.1 & 0.09 \\ 0.09 & 0.1 \end{pmatrix}.$$

The high correlation of 0.9 between the mail and phone ppoulation means is typical of those we have found in analyses of CAHPS data. In addition, the mail and phone response rates at $N = 9$ hypothetical plans are assigned to the 9 distinct pairs of numbers drawn from the set $\{0.35, 0.50, 0.65\}$. We note that in practice the average response rate for mail is typically slightly above 50% and that the phone response rate is often slightly lower than this.

It is clear from Table 1 that SAE enables less costly designs to be used. The cost across the 9 hypothetical plans for full-SAE was 4879 compared to 5764 for the design-based approach and 6666 for hybrid-SAE. The reduction in total cost is due to a smaller total sample size and a smaller proportion of phone surveys. Full-SAE had the lowest cost for every plan whereas the hybrid-SAE proved to be more expensive than the design-based approach for plans with lower mail response rate. The phone sample size constraint was binding for seven plans under the design-based calculation, five for full SAE, and two for hybrid SAE.

Hybrid-SAE yields larger sample sizes than full-SAE largely because of the bias introduced by reweighting the mean outcomes. Because precision is evaluated under the design in the former and the model in the latter, hybrid-SAE is disadvantaged relative to full-SAE. An interesting finding is that hybrid-SAE was sometimes more costly than the design-based approach. It is surprising that the hybrid-SAE approach costs more than the design-based approach under any scenarios. We suspect that this is an artifact of the bias introduced by shrinkage. Conversely the full SAE method reduces the posterior variance through shrinkage but the shrinkage of the posterior means also signifies that comparisons are less significant as well. An alternative metric upon which to constrain the variances would have been to consider the reliability of comparisons between two plans or between a plan and the overall mean; these reflect the contrasts most frequently performed in practice.

To gain further insight into the optimal designs we ran additional simulations at different settings of the design parameters. We found that:

1. Increasing c_{i1} increases the ratio n_{i0}/n_{i1} for all methods.

2. In SAE approaches, decreasing the correlation between the plan means for the mail and phone ratings decreases n_{i0}/n_{i1}.

Table 1: Optimal designs for nine plans under design-based, SAE, and a hybrid SAE-design-based calculation.

		Design-Based			
r_{i0}	r_{i1}	Mail	Phone	Cost	Binding
0.35	0.35	685	446	1576	yes
0.35	0.5	589	349	1287	no
0.35	0.65	550	285	1121	no
0.5	0.35	729	365	1458	yes
0.5	0.5	567	284	1134	yes
0.5	0.65	480	240	960	yes
0.65	0.35	917	321	1559	yes
0.65	0.5	685	240	1165	yes
0.65	0.65	561	196	953	yes
Total		5764	2725	11213	7

		Full SAE			
r_{i0}	r_{i1}	Mail	Phone	Cost	Binding
0.35	0.35	687	346	1380	no
0.35	0.5	592	282	1157	no
0.35	0.65	532	244	1019	no
0.5	0.35	579	290	1158	yes
0.5	0.5	483	242	966	yes
0.5	0.65	431	209	848	no
0.65	0.35	618	216	1051	yes
0.65	0.5	511	179	869	yes
0.65	0.65	446	156	758	yes
Total		4879	2164	9206	5

		Hybrid SAE			
r_{i0}	r_{i1}	Mail	Phone	Cost	Binding
0.35	0.35	992	333	1657	no
0.35	0.5	906	269	1444	no
0.35	0.65	852	231	1313	no
0.5	0.35	767	289	1345	no
0.5	0.5	695	233	1160	no
0.5	0.65	649	199	1047	no
0.65	0.35	686	240	1166	yes
0.65	0.5	585	205	995	yes
0.65	0.65	534	179	892	no
Total		6666	2177	11020	2

Note: The column labeled "binding" indicates whether the phone sample size constraint was binding at the optimal solution.

8 Accounting for the Estimation of β

In this section we outline how the SAE-variance calculations can be extended to accommodate uncertainty in β. If we assume that β has a uniform (improper) prior it follows that

$$\theta_i \mid Y, X \sim N(P_i Y_i + (I - P_i) Y_i^*, V_i^*) \qquad (9)$$

where

$$V_i^* = P_i V_i + (I - P_i) X_i \tilde{V}^{-1} X_i^T (I - P_i)^T$$

$P_i = (V_i^{-1} + \Sigma^{-1})^{-1} V_i^{-1} = \Sigma U_i$, $\tilde{V} = (\sum_i X_i^T U_i X_i)^{-1}$, $Y_i^* = X_i \tilde{V}^{-1} \sum_i X_i^T U_i Y_i$, and $U_i = (\Sigma + V_i)^{-1}$. Note that $(V_i^{-1} + \Sigma^{-1})^{-1} = P_i V_i$ and $I - P_i = U_i \Sigma^{-1}$.

The involvement of $\tilde{V}$ in (9) shows that the estimation of β affects all domains simultaneously. Therefore, optimal sample sizes must be determined simultaneously since a change in the sample size for one plan affects the precision with which β is estimated and thus also of $\eta_i \forall i$. Therefore, the optimization problem involves simultaneous minimization of survey cost across domains.

To find the optimal solution we use the same approach as the β known case but insert an extra step to accommodate uncertainty in β. Each iteration we compute the optimal sample sizes assuming $\tilde{V} = (\sum_i X_i^T U_i X_i)^{-1}$ is a constant. We then update $\tilde{V}$ and re-compute the optimal sample sizes, iterating until convergence. As for the β fixed case, we find the optimal values of n_i by solving the dual problem. A nice feature of this approach is that separate optimization problems are solved for each plan, greatly simplifying the computation.

9 Conclusion

The constraint on phone sample size is an interesting and challenging twist that distinguishes the solution from Neyman allocation.

The solutions we have obtained make sense. Because the predictions from the mail survey lead to more precise estimates of plan means for phone respondents, the marginal benefit of each phone respondent is reduced. Therefore, the optimal number of non-respondents that are followed up by phone should indeed be smaller than under the design-based approach.

In future work we plan to account of the uncertainty in the response rates π_{ij} in the computation of optimal sample size. We also plan to evaluate the designs using methods more closely tied to the problem of comparisons among plans.

Appendix

The Lagrangian is given by:

$$\begin{aligned} L(n_0, n_1, \nu, \omega) &= \sum_i n_{i0} c_{i0} + n_{i1} c_{i1} \\ &+ \sum_i \lambda_i (w_{i0}/n_{i0} + w_{i1}/n_{i1} - v_0) \\ &+ \sum_i \omega_i (n_{i1} - n_{i0}(1 - r_{i0})). \end{aligned} \qquad (10)$$

The first derivatives of the lagrangian in (10) are given by:

$$\frac{\partial}{\partial n_{i0}} L = c_{i0} - \frac{\lambda_i w_{i0}}{n_{i0}^2} - \lambda_i (1 - r_{i0}), \qquad (11)$$

$$\frac{\partial}{\partial n_{i1}} L = c_{i1} - \frac{\lambda_i w_{i1}}{n_{i1}^2} + \lambda_i r_{i0}, \qquad (12)$$

$$\frac{\partial}{\partial \lambda_i} L = \frac{w_{i0}}{n_{i0}} + \frac{w_{i1}}{n_{i1}} - v_0, \qquad (13)$$

and

$$\frac{\partial}{\partial \omega_i} L = n_{i1} - n_{i0}(1 - r_{i0}). \qquad (14)$$

The optimal value of n_i is found by solving $\frac{\partial}{\partial n_{i0}} L = 0$, $\frac{\partial}{\partial n_{i1}} L = 0$, and $\frac{\partial}{\partial \lambda_i} L = 0$. Because (14) is associated with a non-negativity constraint we solve $\frac{\partial}{\partial \omega_i} L \geq 0$. Note that the Karush-Kuhn-Tucker (KKT) conditions (Kuhn and Tucker 1951) also require that $\omega_i (n_{i1} - n_{i0}(1 - r_{i0})) = 0$ and $\omega_i \geq 0$.

Solving (11) and (12) for n_{i0} and n_{i1} respectively we obtain:

$$n_{i0} = \left(\frac{\lambda_i w_{i0}}{c_{i0} - \nu_i (1 - r_{i0})} \right)^{1/2}, \qquad (15)$$

$$n_{i1} = \left(\frac{\lambda_i w_{i1}}{c_{i1} + \nu_i} \right)^{1/2}. \qquad (16)$$

Substituting the expressions for n_{i0} and n_{i1} in (15) and (16) into (13) and $n_{i1} \geq n_{i0}(1 - r_{i0})$, and solving for λ_i and ω_i we obtain

$$\begin{aligned} \lambda_i = v_0^{-2} \Big[& \{ w_{i0} (c_{i0} - \nu_i (1 - r_{i0})) \}^{1/2} \\ & + \{ w_{i1} (c_{i1} + \nu_i) \}^{1/2} \Big]^2 \end{aligned}$$

and

$$\nu_i = \max \left\{ 0, \frac{w_{i1} c_{i0} - w_{i0} c_{i1} (1 - r_{i0})^2}{(1 - r_{i0}) \{ w_{i1} + w_{i0} (1 - r_{i0}) \}} \right\}. \qquad (17)$$

The Lagrange multipler λ_i is essentially a ratio of cost times standard error of estimation to the required precision. Values of $\omega_i > 0$ represent the extent to which the constraint $n_{i1} \leq n_{i0}(1 - r_{i0})$ is slack (i.e., the amount that the phone resource is underused given the mail response rate).

The condition $\nu_i > 0$ occurs if $w_{i1}c_{i0} > w_{i0}c_{i1}(1 - r_{i0})^2$, an event we denote by a_i. Expressing the sample sizes for the $a_i = 0$ and $a_i > 0$ cases in the same form we obtain

$$n_{i0} = \frac{1}{v_0}\left(w_{0i} + \frac{w_{1i}}{t_i}\right),$$

$$n_{i1} = \frac{t_i}{v_0}\left(w_{0i} + \frac{w_{1i}}{t_i}\right),$$

where

$$t_i = (1 - r_{i0})^{a_i}\left(\frac{c_{i0}w_{i1}}{c_{i1}w_{i0}}\right)^{(1-a_i)/2},$$

an alternative expression for t_i given in Section 3.

References

Cochran, W. G. (1977). *Sampling Techniques.* John Wiley & Sons.

Crofton, C., J. S. Lubalin, and C. Darby (1999). The Consumer Assessment of Health Plans Study (CAHPS): Foreword. *Medical Care 37*, MS1–MS9.

Goldstein, E., P. D. Cleary, K. M. Langwell, A. M. Zaslavsky, and A. Heller (2001). Medicare Managed Care CAHPS: A tool for performance improvement. *Health Care Financing Review 22*(3), 101–107.

Kuhn, H. W. and A. W. Tucker (1951). Nonlinear programming. In J. Neyman (Ed.), *Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 481–492. University of California Press.