Optimal Survey Design: A Review

Jeffrey M. Gonzalez* John L. Eltinge†

Abstract
When designing surveys, survey organizations must consider numerous design features that may have a substantial and differential impact on both data quality and survey costs. They must recognize that surveys are inherently multipurpose and that a potentially long list of constraints (e.g., minimum sample sizes for domains) must be satisfied. A typical approach is to optimize an objective function subject to constraints on costs and quality. However, as the list of constraints lengthens and the cost and quality structures become more complex, finding a solution to this optimization problem (i.e., choosing the appropriate set of design features) while satisfying all of the constraints becomes increasingly challenging. This paper reviews the methods by which survey designers have attempted to satisfy multiple constraints while optimizing some function of data quality and survey costs.

Key Words: Constrained optimization, Decision variables, Designed missingness, Multiple objective functions, Stratified multistage sample design, Superpopulation

1. Introduction

When designing new or modifying existing surveys, survey organizations must consider many design features that may have a substantial and differential impact on both data quality and survey costs. Surveys are inherently multipurpose, meaning that there will be many, both anticipated and unanticipated, uses of the collected data. These uses can range from producing estimates of characteristics of the survey's target population to policymakers using the collected data to make policy decisions that may affect many of their constituents. For each distinct use, there are usually specific requirements, or constraints, that must be satisfied so that the data user can have some confidence that the data collected are sufficient for their purpose. As the number of data users increases, the list of constraints that needs to be satisfied may also lengthen. Satisfying the needs of every data user can be challenging because as the number of constraints increases, then finding a solution to this optimization problem becomes increasingly complex. Furthermore, a wide range of additional constraints arise from budgetary and operational considerations.

The purpose of this paper is to review the issues encountered during survey design, identify how survey designers and researchers have addressed these issues, and offer suggestions for extensions of the topics discussed. This review is divided into the following sections. Section 2 provides background information on survey design, by defining key concepts that will be useful to bear in mind throughout this review. Section 3 identifies an effort within the Bureau of Labor Statistics (BLS) motivating this review. Section 4 outlines the basic approach to optimization as it applies to simple survey design problems. Section 5 provides a general optimality framework that may be helpful when exploring issues and trade-offs related to survey design. Section 6 highlights extensions of the basic approach to optimization to

*gonzalez.jeffrey@bls.gov, Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Ave NE, Suite 1950, Washington, DC 20212

†Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Ave NE, Suite 1950, Washington, DC 20212

problems specific to multipurpose, or multiobjective, surveys. Finally, the paper concludes with a discussion and characterization of the trade-offs that often arise during survey design.

2. Survey Design

This section provides background information on survey design. In particular we identify some purposes of surveys, offer a formal definition of survey design, discuss areas of potential conflict that may arise when designing surveys, and highlight what is meant by optimal survey design.

It is important to identify the purposes of a specific survey in advance of electing particular design options. By identifying the purposes of a survey and perhaps prioritizing them, the survey designer can make more informed design decisions about how best to satisfy those objectives. Kish (1988) offers an extensive review of multipurpose sample designs. In his review he provides a hierarchy of six primary purposes of surveys. They are: (1) calculation of diverse statistics; (2) characterization of diverse statistics; (3) collection of multiple variables; (4) multi-subject surveys; (5) continuation of survey operations; and (6) master frames.

There are two points worth mentioning about this list. First, it contains a very diverse set of purposes. Various types of researchers often use surveys to characterize or describe certain populations (related to purposes 1 and 2 above), but they forget that surveys serve a plethora of other purposes. One such purpose is to provide inputs for the development of sampling frames for subsequent surveys (purpose 6 above). The second comment is that practically every survey will attempt to satisfy multiple purposes on this list. In other words, there is overlap among the items on the list. Furthermore, even within a particular purpose, there is a range of possible objectives. For instance, calculation of diverse statistics may not only refer to calculating means or totals but also to computing analytic statistics such as regression coefficients. An added complication is that calculating each type of statistic may require a different set of conditions. The ultimate implication here is that surveys are inherently multipurpose. So, when designing a survey, consideration must be given to the multipurpose nature of the survey.

Given that the purposes of a survey have been identified, the next issue that survey designers must consider is how to construct, or design, a survey to satisfy each purpose. This, however, requires a precise definition of survey design. Kish (1965, Chapter 1) provides a definition of survey design that encompasses a broad range of components, but this actually speaks to the diversity of survey purposes. His definition has two key aspects, which he calls survey objectives and sample design. By survey objectives he means the definition of survey variables, methods of observation, methods of analysis, utilization of results, and desired precision. He refers to the sample design as containing two processes – selection of sample units and estimation from the sample units. It is important to recognize that these two key aspects are not independent of each other. In other words, survey design is a two-way process. In some sense, the survey objectives should determine the sample design, but often sample design issues influence the survey objectives. Thus, it is necessary to have a dialogue among subject matter experts, statisticians, end-users, and anyone else who has knowledge in either the survey objectives or sample design of a particular survey.

When designing a survey to meet some combination of the purposes listed above, conflicts may arise because it is challenging to satisfy every purpose. Kish (1988) identifies ten areas of conflict that may arise when designing multipurpose surveys. They are: (1) sample sizes; (2) relation of biases to sampling errors; (3) allocation of sample among domains; (4) allocation of sample among strata; (5) choice of stratification variables; (6) cluster sizes; (7) measures of size for clusters; (8) retaining sample units; (9) design over time; and, (10) sampling errors. Fortunately not every potential conflict needs to be considered when designing a survey because some conflicts are tied only to one purpose. However, Kish recommends that key considerations should always be given to sample size issues and issues regarding bias ratios because those conflicts tend to be ubiquitous. As a final note, it may appear that these conflicts only pertain to sample design issues, but they are in fact related to the survey objectives. For instance, there may be a particular population subgroup, or domain, that a researcher is interested in describing, so a potential conflict may arise when determining how to obtain information on members of that subgroup given the distribution of its members among various strata.

The ultimate goal in survey design should be to choose the best or optimal design to meet the primary purposes of the survey. For the purposes of this discussion, optimal means choosing the best element from some set of available alternatives. In survey design, there are two perspectives on what is meant by optimal. The first is an "ideal" notion of optimum. This notion refers to the scenario of operating under an unconstrained system. As a simple example, survey research suggests that nonresponse rates tend to be lower with personal visit surveys than they are with either mail or telephone surveys. However, personal visit interviews are generally more expensive than telephone interviews (Groves, *et al.*, 2004, Chapter 5). So, if the survey organization was unconstrained by budgets and if it wanted to optimize, or maximize, the response rate, then it should choose to administer a personal visit survey. Often, however, surveys must operate under constraints, and in particular monetary constraints. In practice, survey design must balance a wide range of factors and there is usually only a finite set of resources with which to conduct a survey. Thus, there is a second notion of optimum which is referred to as a "practical" optimum. This notion of optimum can be thought of as the one that is achieved after specifying certain constraints and conditions on the system.

3. Motivating Application

The impetus for this review is the redesign of the U.S. Consumer Expenditure (CE) Survey. The CE Survey program consists of two surveys, a quarterly personal visit interview survey and a two-week diary. These two surveys combined provide information on the buying habits of American consumers, including data on their expenditures, income, and household characteristics. The survey data are collected for the BLS by the U.S. Census Bureau and provide the basis for revising the weights and associated pricing samples for the Consumer Price Index (CPI), one of the nation's leading economic indicators (BLS *Handbook of Methods*, 2007). With the current surveys, however, there is concern over underreporting of expenditures, declining participation rates, and high respondent burden. To address these concerns an effort, known as the Gemini project, is underway to redesign the CE Survey program. The aim of the project is to provide a detailed road map for a redesigned survey program that will address these concerns and improve data quality.

With a redesign imminent, there is ongoing research about potential ways to modify the surveys. One potential set of methods that is currently being explored is split questionnaire methods (Gonzalez and Eltinge, 2007; 2008; 2009). Briefly defined, split questionnaire methods involve dividing a survey into subsets of questions and administering each of those subsets to subsets of a full, initial sample. The idea is that by reducing respondent burden, if burden is viewed as number of questions asked to a particular respondent, then data quality may improve because the survey will be less burdensome and a less burdened respondent may be more motivated to provide more accurate responses (Gonzalez and Eltinge, 2007).

It is worth clarifying that split questionnaire methods or adaptations of them are only one of the many potential ways that the CE Survey program is considering during the redesign effort. In fact, other potential methods currently being investigated are the use of global questions, altering the recall period, and changing the interview structure (http://www.bls.gov/cex/geminiproject.htm). Regardless of the way the CE survey program is redesigned, it is essential to have a framework for studying the extent to which a particular redesign option is optimal. This framework will enable the CE Survey program to determine whether the redesign option maximizes some function of quality or other important characteristic. Therefore, one of the goals of this review is to begin developing that framework. Specific details on the proposed framework can be found in Section 5.

4. Review of Basic Approach to Optimization

Expressing a survey design problem in mathematical notation is the key to selecting a particular design with specified properties. This is because explicit formulas representing primary survey purposes can easily be optimized subject to constraints and/or requirements of data users. There are four primary components of an optimization problem as it pertains to survey design. They are: (1) the objective function; (2) decision variables; (3) parameters; and (4) constraints. In this section, we provide both definitions and examples of each component. It is important to bear in mind that the listing of examples is by no means exhaustive.

The first component of an optimization problem is the objective function. The objective function is a function of one or more variables to be optimized. We can think of optimization in terms of maximization or minimization. In simple problems, examples of an objective function include the theoretical sampling variance of a prespecified estimator of a population quantity of interest. Below we have examples of two standard sampling variance formulas for, respectively, stratified random sampling and stratified unequal-probability sampling:

$$V(\hat{\bar{y}}) = \frac{1}{N^2} \sum_h N_h (N_h - n_h) \frac{S_h^2}{n_h} \tag{1}$$

$$V(\hat{\bar{y}}_k) = \frac{1}{N^2} \sum_h \sum_{U_h} \left[\left(\frac{1-\pi_i}{\pi} \right) + w_i \left(\frac{1-p_{hi}}{p_{hi}} \right) \right] (y_{hik} - \bar{Y}_k)^2 \tag{2}$$

The above formulas can be extended in many ways. For instance, one can derive formulae that account for multivariate y-vectors (to account for several characteristics of interest), multiple stages and/or phases of sampling (equation [2] can be used for two phases of sampling), and different estimators (such as estimators for totals and regression coefficients).

The second component of an optimization problem is the set of decision variables. These are the quantities that are adjusted in order to find a solution to the optimization problem. Essentially, these serve as the primary outputs, or what the survey designer is most interested in. Examples of decision variables may include full or stratum-specific sample sizes, denoted by n and n_h, respectively, first-order inclusion probabilities, π_i, and subsampling probabilities for two-phase designs, p_{hi}.

The next component of the problem is a set of fixed inputs that are treated as constants. These are known as the parameters. Examples of parameters can be identified using the objective function examples given above. For instance, in equation (1) we have population stratum variances and stratum-specific population sizes, denoted by S_h^2 and N_h, respectively. Examples from equation (2) include population size, N, and values of specific variables on population units, y_{hik}. Another important example of parameters not contained in either equation are cost components. These include the cost of observing a specific unit in a particular stratum. These can be denoted as c_h.

The final component is the set of constraints. These are the restrictions on the decision variables or combinations of them. Examples of constraints may include specifying interviewer workloads, specifically assigning a certain number of cases to each interviewer. Another constraint may be to observe a minimum number of sample units in a particular stratum or domain. Equation (3) displays this specific constraint.

$$n_h \geq n_{min} \tag{3}$$

As mentioned earlier, survey operations are often constrained by budgets, so another example of a constraint is a cost constraint. Equation (4) identifies a simple linear cost constraint in which the overall survey cost, denoted by C, is composed of a fixed cost component, C_0, as well as a variable component that depends on the number of observations made in each stratum, per Cochran (1977, Chapter 5).

$$C = C_0 + \sum_h c_h n_h \tag{4}$$

We now illustrate the basic approach to optimization using some of the specific examples above. A simple variance-cost optimization is as follows: we would like to determine the set of stratum sample sizes $\{n_h\}$ that minimize (1) subject to the constraint identified in (4). Using either an application of the Cauchy-Schwarz inequality (Cochran, 1977, Chapter 5) or Lagrange multipliers (Varberg and Purcell, 1997, Section 15.9), one can easily find that the solution is:

$$n_h = n \frac{N_h S_h / \sqrt{c_h}}{\sum_h N_h S_h / \sqrt{c_h}} \tag{5}$$

In other words, the set of $\{n_h\}$ identified in (5) will minimize (1) subject to (4). The implication of allocating sample based on (5) is that we sample more from strata that are heterogeneous (have large S_h) and less from strata that are expensive (have high c_h).

The above problem and more complex problems, such as those for multipurpose surveys, can be expressed in general mathematical terms. For example, if we let $f : \Re^d \to \Re$ denote the objective function and $X \subseteq \Re^d$, then we wish to find some element of X, namely x, that maximizes f. We should mention that X is the space of design decisions and d is the dimension of the decision variables vector. In the simple variance-cost optimization problem given above, d is equal to the number of strata.

It is worth noting that these general mathematical terms can also be extended in various ways. For instance, note that the function f is a mapping from $\Re^d$ to $\Re$. This means that when evaluated f takes on a scalar quantity and design decisions are made according to higher (or lower) univariate values. There may be situations when reducing the objective function evaluation to a scalar quantity is problematic; thus, we can represent the optimization problem even more generally by relaxing this requirement.

Although not identified as a main component, the method used to find the optimum is a critical element of the optimization problem. As demonstrated by the example above, simple problems can usually be solved using Lagrange multipliers or applications of the Cauchy-Schwarz inequality (Cochran, 1977, Chapter 5). However, for more complex situations (e.g., multipurpose surveys) more sophisticated techniques are usually needed. Mathematical programming methods, are useful for solving these complex problems. One specific example can be found in Leaver, *et al.* (1996). The authors used nonlinear programming techniques to determine the optimal allocation of data collection resources that would minimize the sampling variance of price change, subject to various budgetary and operational constraints. Their research provided an explicit example of a very complex optimization problem that was solved with the assistance of computers. However, the key to using computational power to solve this problem and others like it is providing accurate inputs into the appropriate interface.

5. General Optimality Framework

While the basic approach to optimization is helpful when setting up survey design problems for computational exercises, a more general optimality framework is useful for exploring issues related to optimal survey design. Specifically, the framework helps survey designers understand the relationships among design decisions, survey data quality, and the utility of statistical products (e.g., official estimates of means, totals, or other quantities) from the collected survey data. This framework is also helpful when exploring general cost and quality trade-offs related to survey programs and their stakeholders. For the purposes of this discussion, a stakeholder is essentially any entity (e.g., person or organization) with a vested interest in the survey program and/or any products subsequently produced from the collected survey data. Specific examples of stakeholders are policymakers and academic researchers.

Key components of this framework can be extracted and adapted from various references on optimal design and statistical decision theory (Fedorov, 1972; Silvey, 1980; Berger, 1980). To develop this optimality framework, we provide the following notation. First, let $\mathcal{D}$ be the decision, or design, space; D denote the selected design feature (e.g., random mechanism for sampling); and, d be the realization of the specific design feature (e.g., sample). In addition, let Q be the optimality criterion (e.g., mean squared error of a survey statistic or one of the six dimensions of data quality outlined in Brackstone [1999]) and U be a utility function representing a stakeholder's relative satisfaction with the design.

We also identify X as a vector of observable auxiliary information. For the current discussion, we partition $X = (X_R, X_B, X_C)$ where X_R is a set of resources with which to conduct a survey (e.g., existing survey organization infrastructure, interviewing staff); X_B are the bounds or constraints (e.g., on data collection budgets); and, X_C is the cost structure (e.g., per unit interview costs). We also allow for the possibility of other factors that are neither observable nor directly controllable in real-time and we denote this vector as Z. An example of this may be changes to the underlying survey environment. A specific example relevant to the CE survey program is that of new products (e.g., iPads or Amazon Kindles). Expenditure information about new products is difficult, if not impossible, to obtain after survey design decisions have been made and resources have been allocated. Thus, these types of changes in the underlying survey environment will have an impact on the optimality criteria and subsequent measure of utility for each stakeholder. As a final note, auxiliary information in both X and Z can either be fixed or random and known or unknown throughout the entire survey process. For example, survey organizations may experience interviewer turnover during data collection or per unit interview costs may increase due to extreme changes in gasoline prices.

Using the above notation, the optimality criterion, Q, can be expressed as the following function.

$$Q = Q(D, X, Z, \gamma) \tag{6}$$

Note that in equation (6) we also have a vector of parameters, denoted by γ. This vector contains parameters that are unknown and may have an impact on the optimality criterion. Specific examples of these may include superpopulation model parameters or design effect parameters associated with the class $\mathcal{D}$ of designs.

Finally, we express the stakeholder's utility function, equation (7), where β is a vector of parameters representing underlying perceptions of needs of individual stakeholders.

$$U = U(Q, \beta) \tag{7}$$

Given this representation of a stakeholder's utility, it is clear that for the same optimality criteria, individual stakeholders' perception of value, or utility, may still vary across stakeholder, due to differences in β. Said differently, while survey designers may make design decisions using one criteriom (e.g., mean squared error of a particular statistic), the value of the survey design may be high for one stakeholder, but quite low for another because that stakeholder has relatively low interest in the particular statistic. For example, some stakeholders' perception of data utility will depend heavily on timeliness, while for others timeliness may be much less important than item-level missingess rates. In addition, if groups of stakeholders with relatively similar utility functions are identified, then using a relevant optimality criterion for those groups, may yield a survey design with greater utility for more stakeholders. Thus, having a clear understanding of various stakeholders' utility functions is an important component of the survey design process because this is ultimately related to choosing an appropriate objective function to optimize.

6. Extensions of Basic Approach to Optimization

Now that we have provided both a foundation for solving survey optimization problems in simple settings and a general optimality framework for discussing issues often encountered in survey design, in this section, we identify ways that survey designers have extended the basic approach to optimization. Specifically, we first identify extensions to applications involving multiple estimands and multiple utility functions. Second, we provide an alternate perspective on survey design and identify specific problems within this perspective that have warranted consideration of optimization methods.

6.1 Multiple Estimands and Multiple Utility Functions

In the example provided in Section 4, we determined the sample allocation among strata based on estimating one unknown population parameter, $\bar{y}_k$. It is worth noting, however, that univariate methods are not necessarily optimal when designing multivariate surveys, or surveys in which there are multiple estimands of interest. For example, strata for stratified sampling are often formed based on known auxiliary information, say H, that is thought to be correlated with the key survey variables. Estimators based on these key survey variables would likely be very efficient due to their high correlation with H. If, however, there are other survey variables of interest that are only weakly correlated with H then stratified sampling based on strata formed by H might not be as efficient for these variables. So, several authors have developed methods and provided criteria for deciding among sample design alternatives for multipurpose, or multivariate, surveys.

Many of the methods that survey designers and statisticians have proposed effectively amount to reducing the evaluated objective function (e.g., sampling variances or coefficients of variation) of several key survey variables to a scalar quantity. For instance, Holmberg (2002) identified three criteria that can be used in multivariate surveys: (1) sum of the variances of estimators under consideration; (2) sum of squared coefficients of variation of those estimators; and, (3) sum of relative losses of efficiency. In addition Kozak (2006) presents and compares five sample allocation methods (one of which is proportional allocation – to serve as a baseline for comparison to the other methods) that could be useful when designing and planning multiparameter surveys. Details on the specific methods being compared can be found in the article. In this simulation study, each method that took into account the multivariate nature of the data outperformed proportional allocation. It is worth noting that proportional allocation does not take into account any of the characteristics of the population other than the population size of each stratum. The primary conclusion of this research was that when choosing an allocation method for a multiparameter survey, survey designers should choose the method that takes into account the multivariate structure of the data being collected.

The criteria reviewed in the previous paragraph may also be modified by assigning importance weights to certain parameters or characteristics of interest (Kish, 1988; Kozak, 2006). For multipurpose surveys, one purpose might be deemed more important than another, so a higher weight can be assigned to that purpose. The the entire set of weights $\{w_k\}$ can then be incorporated into the objective function. As an example of how importance weights can be assigned, consider the following. Suppose we wanted to find the set of stratum sample sizes $\{n_h\}$ to minimize an objective function that is equal to the sum of variances of the K means of interest. This is given by equation (8).

$$\Phi_0 = \sum_{k=1}^{K} var(\hat{\bar{y}}_k) \tag{8}$$

If some of the means are of more interest than the others, then we can assign higher weights to the variances of those means. So, the objective function can be modified by including a set of weights $\{w_k\}$ in the following way and optimization of Φ_1 can occur in the standard way.

$$\Phi_1 = \sum_{k=1}^{K} w_k var(\hat{\bar{y}}_k). \tag{9}$$

Day (2009) also investigated the problem of allocating sample among strata when there are multiple parameters of interest. One of the main differences between this and prior research is that previous studies have assumed that the strata were fixed in advance of determining the sample allocation to each stratum. One reason that a survey designer might not want to fix stratum boundaries in advance of allocating sample is that there may be characteristics of interest that are only weakly correlated with the stratification variable. So, stratified sampling for those characteristics would not be as efficient as it would be if the correlation was higher. Day discussed the use of evolutionary algorithms, a type of metaheuristic algorithm, for the situation of simultaneously determining stratum boundaries and sample allocation among strata. Briefly defined, metaheuristic algorithms are computational techniques that solve the optimization problem iteratively by searching and comparing candidate solutions across some measure of quality or an evaluated objective function. The idea here would be to actively search the space of possible solutions instead of evaluating derivatives to find the extremum of the objective function. In particular, though, evolutionary algorithms employ a biological evolution model for computing. Specific details on these biological models can be found in various references on evolutionary computation (De Jong, 2006).

According to Day (2009), one of the advantages of using an evolutionary algorithm is that a smaller sample size may be used to reach coefficient of variation targets when stratum boundaries and allocations are simultaneously adjusted. His research provides justification that survey design issues can be addressed using these computational techniques, however, further research is needed to understand how to adapt these algorithms to individual problems. In fact, the author lists some areas for future study, e.g., methods for sampling from various distributions. It is worth mentioning that a potential drawback of these methods is that the algorithms have to be tailored to individual problems. The possible lack of generality of these algorithms to a large class of design problems may make survey designers opt for more standard techniques. If, however, an evolutionary algorithm is used, to the extent possible, the resulting outputs should be compared against existing, or more universally accepted, methods using simulations and sensitivity analyses.

Even though sophisticated computational techniques are available to solve multivariate design problems, it may still be desirable to opt for a more simplistic approach. Rahim and Jocelyn (1994) proposed an aggregate measure of the variabilities of all estimates in terms of a distance function of the coefficients of variation. They advocate for this approach because when survey cost is preassigned, having too many individual variance constraints might result in a survey design that exceeds budgetary constraints. When compared to convex programming methods, they concluded that as long as the aggregate measure of variabilities did not exceed its preassigned limit, then there was little point in being overly concerned that each variance constraint was satisfied since their method performs well against convex programming. So, they conclude that taking this approach is promising not only for its simplicity and potential to lower survey costs, but also because there are few violations of the individual sampling variance constraints.

6.2 Unconditional and Conditional Approaches to Sample Selection and Assignment of Collection Methods

Taking a different perspective, one can view optimal design as a process of determining the best "treatment" to administer to each sample member. For the purposes of this discussion, treatment is taken to be very general with examples including, but not limited to, more intensive follow-up/callback procedures for initial nonrespondents, mode of data collection or enumeration, and incentives. Furthermore, a few broad classes of survey techniques where optimization methods might be useful and applicable are responsive designs (Heeringa and Groves, 2004; Groves and Heeringa, 2006), multiphase designs (Sarndal, *et al.*, 1992, Chapter 9), and partitioned, or split questionnaire, designs (Raghunathan and Grizzle, 1995; Gonzalez and Eltinge, 2007; 2008).

An early example of optimization methods being considered to determine treatment assignment to sample members can be found in Wolter (1978). He examines the question of how the entire sample should be optimally allocated to various modes knowing in advance that each mode has different error distributions and cost properties (e.g., per unit costs). In addition, he proposes a composite estimation method to account for sample members being enumerated by a different mode. This method is effectively a weighted sum of the estimates based on data collected according to each mode. His work provides a rationale for assigning different treatments to sample members and a solution for combining the different enumeration methods, but use of these methods requires good prior information about the measurement error distributions and the per unit costs for enumerating using each mode.

Specific treatments to administer to sample members and broad classes of survey techniques can be considered in tandem when exploring optimization methods. For instance, responsive design methods may involve assigning treatments to respondents during the course of data collection. Briefly defined, the term responsive design (Heeringa and Groves, 2004) is used to describe the act of making mid-course decisions and survey design changes based on accumulating process and survey data. Decisions are meant to improve the error properties of the resulting statistics. Although their formal definition of responsive design has five components, the one regarding the decision rule, in which the survey operator actively changes the survey design features in subsequent phases (based on information collected in previous phases), might benefit most from the consideration of optimization methods.

One example of treatments and responsive design methods being considered jointly can be found in Gonzalez and Eltinge (2008). They discuss an adaptive, or responsive, assignment of split questionnaires to sample members participating in a panel survey. In their study, the "treatment" of interest was a split questionnaire, i.e., a shortened version of the original questionnaire. Furthermore, it was adaptive because determination of which split questionnaire to administer to a respondent was based on information collected in the initial interview. Their evaluation of various allocation methods of the split questionnaires to sample members primarily involved criteria that are frequently used in the optimal experimental design literature. These included the following variance-minimizing criteria – A-optimality and D-optimality (Silvey, 1980). Using these two criteria is equivalent to using the smallest trace and smallest determinant of the covariance matrices, respectively. It is worth mentioning that for a covariance matrix, the trace is equal to the sum of variances and the determinant is equal to the generalized variance. Details on which allocation method was deemed superior by the authors can be found in the article. The primary conclusion, however, to draw from their research is that "treatment" assignment, responsive design techniques (and perhaps, more generally, other broad classes of survey techniques), and optimal design criteria can all be used simultaneously to address various issues in survey design.

Another example can be found in Elliot *et al.* (2000). They discuss optimal design procedures for subsampling of callbacks to improve survey efficiency. The broad class of survey techniques that is applicable to their research can include either multiphase designs or responsive designs due to the subsampling and multiple phases of data collection. Primarily motivated by cost savings, the authors examined whether efficiency of data collection could be increased by subsampling a random proportion of the initial sample from a prespecified callback. In this application the "treatment" would be whether or not the sample unit was "called back." Their analysis suggests that randomly dropping a subset of sample units to call back can save resources whenever (1) the per callback or per interview cost is increasing, or (2) the probability of a successful interview attempt is decreasing. They caution that increased variance of the variable of interest among late callbacks or late relative between-callback stratum variability reduces the effectiveness of subsampling.

Additional research endeavors involving "treatment" assignment can be found in Demnati, *et al.* (2007) and O'Malley and Zaslavsky (2007). The first discusses the allocation of sample drawn from multiple frames while the second focuses on nonrespondent subsampling for follow-up.

7. Discussion

As previously noted, survey design will require trade-off decisions regarding the cost and quality properties of the survey. This is because as the number of purposes increases and/or the list of constraints lengthens, then finding an optimal solution to meet every purpose while satisfying each constraint becomes increasingly complex. To deal with this situation, one approach might be to eliminate one (or more) purpose(s) and then solve the problem. However, ignoring certain purposes can result in substantial losses of efficiency for that purpose and sometimes that loss can be more than anticipated. The same can be said for constraints. Having an overly constrained system may make finding a feasible solution incredibly challenging, but reducing the number of constraints might actually affect whether or not a particular purpose of the survey is met. Similar complications may arise if we choose designs that depart slightly from the optimum. So, before finalizing a design one should examine the impact of ignoring certain purposes, relaxing or reducing constraints, and choosing designs that deviate slightly from optimal via simulation studies and sensitivity analyses. As an example, Shimizu and Cai (2008) conducted research on the potential impact of using particular sample allocations when investigating alternative sample designs for a redesigned 2010 National Hospital Discharge Survey. They conducted various simulations to compare (on the basis of relative standard errors) sample allocations to strata for samples of different sizes that were optimized using either Neyman allocation of nonlinear programming.

There are two additional topics that warrant discussion. The first is data quality. In the simple example presented in Section 4, we illustrated the following optimization problem: the theoretical stratified sampling variance of the population estimator for the mean was minimized subject to constraints on survey costs. An implicit assumption of using the sampling variance of one population estimator as the objective function is that we are only focused on one aspect of data quality, namely sampling error. It is important to recognize that are other components of data quality.

There are essentially two paradigms for data quality. The first is referred to as the Total Survey Error (TSE) paradigm (Groves, *et al.*, 2004) and the second is known as the Total Quality Management (TQM) paradigm (Biemer and Lyberg, 2003). The former focuses on how at every stage of the survey process both systematic and variable errors can arise. The primary focus of this paradigm tends to be the accuracy of the survey estimates and consists of the following errors or error sources: coverage, sampling, nonresponse, measurement, processing, and post-survey adjustment error and construct validity. On the other hand, the TQM perspective on data quality includes accuracy, and all types of errors, as a single dimension of data quality, but also incorporates dimensions relevant to a data user's perspective namely, relevance, timeliness, coherence, interpretability, and accessibility (Brackstone, 1999). The main point is that it might be worth incorporating additional ideas from these two perspectives on data quality into various components of the optimization problem.

The final point deals with the inputs, or set of parameters, into the optimization problem. These parameters will directly affect the outcome of the survey design; thus, it is important to obtain good prior information that can be used as inputs into the optimization problem. Cochran (1977, Chapter 4) identifies four ways in which a survey designer can obtain prior information necessary for the design. They are: (1) taking a smaller random sample from the population and estimating the desired quantities; (2) the results of a pilot study; (3) previous administrations of the survey to the same or similar population; and, (4) guesswork about the structure of the population, aided by mathematical models. One source of prior information that is absent from this list is data users. Often, data users' have a wealth of information which may provide survey designers with additional insight into the survey optimization problem. Not only may the data users be able to provide values to the parameters, but they may also be able to offer suggestions for evaluative criteria to be used in the objective function and suggest constraints that are needed on the survey system. Therefore, whatever form the survey optimization problem has, it is imperative to acknowledge that survey data users are key components of the problem as well.

8. Acknowledgments

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the U.S. Bureau of Labor Statistics. The authors would also like to thank Ashley Bowers, Charles Day, Jennifer Edgar, Scott Fricker, Karen Goldenberg, Frauke Kreuter, Jay Ryan, Adam Safir, Lucilla Tan, and Rick Valliant for their contributions to this paper.

REFERENCES

Bellhouse, D. R. (1984). A review of optimal designs in survey sampling. *The Canadian Journal of Statistics*, 53-65.

Berger, J. O. (1980). *Statistical Decision Theory: Foundations, Concepts, and Methods.* New York: Springer-Verlag.

Biemer, P. P. and Lyberg, L. E. (2003). *Introduction to Survey Quality.* New York: Wiley.

Brackstone, G. (1999). Managing data quality in a statistical agency, *Survey Methodology*, 25, 139–49.

Bureau of Labor Statistics, U.S. Department of Labor, *Handbook of Methods*, Chapter 16, April 2007 edition, Consumer Expenditures and Income. http://www.bls.gov/opub/hom/pdf/homch16.pdf (visited May 10, 2010).

Chatterjee, S. (1972). A study of optimum allocation in multivariate stratified surveys. *Skandinavisk Actuarietidskrift*, **55**, 73-80.

Cochran, W. G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *The Annals of Mathematical Statistics*, **17**, 164-77.

Cochran, W. G. (1977). *Sampling Techniques*, Third Edition. New York: John Wiley and Sons.

Dalenius, T. (1953). The multivariate sample problem. *Skandinavisk Actuarietidskrift*, **36**, 92-102.

Day, C. D. (2006). An application of genetic algorithms to multivariate optimal allocation in stratified sample designs. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 2931–7.

Day, C. D. (2009). Evolutionary algorithms for optimal sample design. *A paper presented at the 2009 Federal Committee on Statistical Methodology Conference, Washington, DC.*

De Jong, K. A. (2006). *Evolutionary Computation: A Unified Approach.* MIT Press, Boston, MA.

Demnati A., Rao, J. N. K., Hidiroglou, M. A., and Tambay, J. -L. (2007). On the allocation and estimation for dual frame survey data. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 2938–45.

Fedorov, V. V. (1972). *Theory of Optimal Experiments.* New York: Academic Press.

Gentle, J. E., Narula, S., and Valliant, R. (1997). "Multicriteria optimization in sampling design," in *Statistics of Quality*: Dedicated to Don Owen, eds. S. Ghosh, W. Schucany, and T. Smith, New York: Marcel Dekker, 411-25.

Gonzalez, J. M. and Eltinge, J. L. (2007). Multiple matrix sampling: A review. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 3069–75.

Gonzalez, J. M. and Eltinge, J. L. (2008). Adaptive matrix sampling for the Consumer Expenditure Interview Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 2081–8.

Gonzalez, J. M. and Eltinge, J. L. (2009). Imputation methods for adaptive matrix sampling. *A poster presented at the 2009 Joint Statistical Meetings, Washington, DC.*

Gren, J. (1964). Some methods of sample allocation in multiparameter stratified sampling. *Przeglad Statystyczny*, **11**, 361-9 (in Polish).

Gren, J. (1966). Some application of non-linear programming in sampling method. *Przeglad Statystyczny*, **13**, 203-17 (in Polish).

Groves, R. M. (1989). *Survey Errors and Survey Costs.* New Jersey: John Wiley and Sons.

Groves, R. M., Fowler J., Couper M. P., Lepkowski J. M., Singer E., and Rourangeau R. (2004). *Survey Methodology*, New York: Wiley.

Groves, R. M. and Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey error and costs. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, **169**(3), 439–57.

Hartley, H. O. (1965). Multiple purpose optimum allocation in stratified sampling. *Proceedings of the Social Statistics Section, American Statistical Association*, 258-61.

Heeringa, S. G. and Groves, R. M. (2004). Responsive design for household surveys. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 3644–51.

Holmberg, A. (2002). A multiparameter perspective on the choice of sampling design in surveys. *Statistics in Transition*, **5**, 969-94.

Kish, L. (1965). *Survey Sampling.* New York: John Wiley and Sons.

Kish, L. (1988). Multipurpose sample designs. *Survey Methodology*, **14**(1), 19-32.

Kokan, A. R. and Khan, S. (1967). Optimum allocation in multivariate surveys: An analytical solution. *Journal of the Royal Statistical Society B*, **29**, 115-25.

Kozak, M. (2006). On sample allocation in multivariate surveys. *Communications in Statistics - Simulation and Computation*, **35**, 901-10.

Lasdon, L. and Smith, S. (1992). Solving sparse nonlinear programs using GRG. *INFORMS Journal on Computing*, **4**, 2-15.

Lasdon, L., Waren, A., Jain, A., and Ratner, M. (1978). Design and testing of a generalized reduced gradient code for nonlinear programming. *ACM Transactions on Mathematical Software*, **4**, 34-50.

Leaver, S. G., Johnson, W. H., Baskin, R., Scarlett, S., and Morse, R. (1996). Commodities and services sample redesign for the 1998 Consumer Price Index revision. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 239–44.

Nenov, I. and Fylstra, D. (2003). Interval methods for accelerated global search in the Microsoft Excel Solver. *Reliable Computing*, **9**, 143-59.

O'Malley, A. J. and Zaslavsky, A. M. (2007). Optimal survey design when nonrespondents are subsampled for followup. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 3268–74.

Powell, S. G. and Baker, K. R. (2003). *The Art of Modeling with Spreadsheets: Management Science, Spreadsheet Engineering, and Modeling Craft.* New York: John Wiley and Sons.

Raghunathan, T. E. and Grizzle, J. E. (1995). A split questionnaire survey design. *Journal of the American Statistical Association*, **90**, 54–63.

Rahim, M. A. (1995). Use of distance function to optimize sample allocation in multivariate surveys: A new perspective. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 365-9.

Rahim, M. A. and Jocelyn, W. (1994). Sample allocation in multivariate stratified design: An alternative to convex programming. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 689-92.

Sarndal, C. E., Swennsson, B., and Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Scott, A. J. and Smith, T. M. F. (1975). Minimax designs for sample surveys. *Biometrika*, **62**(2), 353-7.

Shimizu, I. and Cai, R. (2008). Some research on sampling design alternatives for a redesigned 2010 National Hospital Discharge Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1460-7.

Silvey, S. D. (1980). *Optimal Design*. New York: Chapman and Hall.

Smith, T. M. F. (2001). Biometrika centenary: Sample surveys. *Biometrika*, **88**(1), 167–94.

Spencer, B. D. (1982). Technical issues in allocation formula design. *Public Administration Review*, **42**(6), 524-9.

Valliant, R. and Gentle, J. E. (1997). An application of mathematical programming to sample allocation. *Computational Statistics and Data Analysis*, **25**, 337-60.

Varberg, D. and Purcell, E. J. (1997). *Calculus*, Seventh Edition. Prentice Hall: New Jersey.

Wywial, J. (1988). Minimizing the spectral radius of means of vector from sample variance-covariance matrix sample allocation between strata. *Prace Naukowe Akademii Ekonomicznej we Wroclawiu*, **404**, 223-35 (in Polish).

Williams, S., Zhao, Z., and Potter, F. (2007). Survey designs to optimize efficiency for multiple objectives: methods and applications. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 3422-7.

Wolter, K. M. (1978). Optimum sample allocation to different modes of enumeration. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 414-9.