Journal of Statistical Planning and Inference 48 (1995) 229–240

journal of statistical planning and inference

Selecting multiple-objective fixed-cost sample designs using an admissibility criterion

Donald Malec*

National Center for Health Statistics, Centers for Disease Control and Prevention, 6525 Belcrest Road, Hyattsville, MD 20782, USA

Received 23 October 1992; revised 31 October 1994

Abstract

An admissibility criterion is used to identify a set of designs useful for multiple-objective surveys. This admissibility criterion is defined over the range of survey objectives, unlike the usual situation where an admissibility criterion is defined over the parameter space. Assuming a fixed total cost, known (or estimated) finite population characteristics and specified estimators, the admissibility criterion is shown to apply to a wide range of design problems. An example using PPS sampling is discussed extensively. Here, one must choose the selection probabilities when there are multiple objectives. Other suggestions proposed in the literature for multiple-objective fixed-cost designs are reviewed with the aim of identifying both methods which produce admissible designs and methods which may produce inadmissible, hence inferior, designs. Lastly, suggestions for picking a final design are made. It is recommended that a minimax design or minimax subset design be chosen from the admissible set.

AMS Subject Classification: Primary 62D05; secondary 62C15, 62C20

Keywords: Compromise sample design; Multivariate sample design

1. Introduction

Most large-scale surveys serve many purposes. Primarily, this is due to the high cost of initiating and fielding a survey and the relatively low cost of adding extra questions to a survey instrument. As a result, a typical survey instrument may contain hundreds of items to be obtained for each sampled unit. The ideal sample design would provide precise estimates of finite population parameters for each of these items. In addition, the survey may also be used to provide estimates for selected subpopulations (e.g., race

* E-mail: djm6@nch09a.em.cdc.gov.

SSDI 0378-3758(94)00153-7

by sex groups). Again, a good sample design will provide precise estimates of population values for each subpopulation and, also, each item of interest. Often, the survey planner does not have a fixed level of precision in mind for each survey objective but the funds available for the survey are fixed. In this situation, the survey planner would like to see what levels of precision can be obtained for a given total cost. The selected fixed-cost design should be efficient, in the sense that no other design with the same total cost has uniformly better precision for each variable/subpopulation combination.

Selecting an optimal sample design to estimate a finite population parameter corresponding to a single variable has been well formulated for many situations. For example, the optimal allocation of a stratified, multistage sample can be found in sampling-theory textbooks such as Cochran (1977). On the other hand, the type of design discussed here, i.e., the multipurpose fixed-cost sample design has not been extensively developed. Although design recommendations for a number of specific multipurpose designs have been made (Bankier, 1988; Chatterjee, 1967; Folks and Antle, 1965; Yates, 1960), there are no general recommendations available.

One reason for the delay in providing methodology for the choice of a fixed-cost, multiple-objective design is that one design will never be optimal for all objectives (if these objectives are different from each other). Another reason for this lack of development may be the large dimensionality of the problem. For example, estimates from the National Health Interview Survey are routinely published for more than 70 variables. For selected variables, estimates are also provided for over 22 domains, consisting of race, age, sex and their cross classifications (NCHS, 1988). Besides the use of published estimates, various types of data analysis, modeling and hypothesis testing are performed by users of the NHIS public use tapes. Designing a survey that performs well for all of these objectives is typically infeasible. In fact, it may not even be feasible to check how well a design performs for each objective.

This paper presents a normative method of multipurpose sample design based on the application of decision theory. In Section 2, the property of admissibility is used to select a subclass of designs where only the members have the desired properties. This is accomplished using an admissibility criterion which has the effect that the separate survey objectives are combined into a single objective function. As will be seen, the admissibility criterion corresponds to the criterion of minimizing the weighted average of individual risk functions, a multiple objective design criterion (with unknown optimality properties) already in use. In Section 3, the problem of selecting a single design from the admissible class is addressed. Previously proposed methods for choosing a survey design for multiple objectives are reviewed and it is shown that many of the resulting designs are admissible. A method which uses a minimax-type criterion is suggested and illustrated.

Throughout this development, finite population design parameters (i.e., coefficients of variation (CVs) or variances) are assumed known. In the design of repeated surveys, good estimates of CVs are often available. Also, for each design, it is assumed that a specific estimator will be used.

A different multipurpose sample design problem, minimal cost design, will not be addressed here. For a variety of designs, the minimal cost, fixed precision, multiple-objective design can be obtained using nonlinear programming techniques. See, for example, Kokan (1963) and Huddleston et al. (1970). Unlike the design problems addressed in this paper, this type of design selection requires that for each objective the desired precision for the corresponding estimator be specified in advance.

2. Admissible designs

In this section, the notation and principle result are specified. Then, examples are used to illustrate both the single purpose and multipurpose design problems.

Let $b \in B$ denote one of a set of K survey objectives. For a specific objective, b, let X_b denote the corresponding finite population parameter of interest. Given a fixed total sampling cost C, one design $d \in D$ will be selected from the set of candidate designs D. For each d, let $\hat{X}_b(d)$ denote the estimator of X_b based on design d. The form of the estimator must be specified so that it can be constructed once the design is known and the sample is selected (e.g., a simple expansion or ratio estimator).

2.1. Optimal single purpose design

In a single purpose design $K = 1$ and the optimal design can be obtained from the two steps.

(a) Specify a loss function $L(\hat{X}(d), X)$. Typically, squared error loss $(\hat{X}(d) - X)^2$, or relative squared error loss $((\hat{X}(d) - X)/X)^2$ are used in survey sampling. However, loss functions that express the error of a hypothesis test or the loss associated with selecting a particular sample design for a specified data analysis may be used.

(b) Subject to any constraints (e.g., fixed costs), choose the design $d \in D$, which minimizes the risk function $R(d) = E\{L(\hat{X}(d), X)\}$. Typically, the expectation is over the sampling distribution induced by the design d.

In most applications, only one design from the set of candidate designs is optimal (see Cochran (1977, Sections 5.5 and 10.6) for examples of optimal single purpose designs).

2.2. Optimal multipurpose design

First, a loss function $L_b(\hat{X}_b(d), X_b)$ must be specified for each $b \in B$. The loss function for each objective can be determined as if the survey was single purpose. Then, based on the following admissibility criterion, a subset of optimal designs can be obtained.

Definition. Let $R(b, d)$ denote the risk function $E\{L_b(\hat{X}_b(d), X_b)\}$. A design $d \in D$ is admissible for the objective of estimating each X_b, $b \in B$ (based on estimators $\hat{X}_b(d)$)

if, and only if, there is no $d_0 \in D$ which is better than d. A design d_0 is better than d if $R(b, d_0) \leqslant R(b, d)$ for all $b \in B$, with strict inequality for at least one b.

Note that this admissibility criterion is defined over the range of survey objectives, unlike the usual situation where admissibility is defined over the parameter space.

Clearly, admissible designs are preferable and they can be identified using three theorems in decision theory (see, e.g., Berger (1985, Theorems 4.7, 4.8 and 5.10)). The result that is needed is: Each design d in the class of designs that minimizes

$$r(\lambda, d) = \sum_b \lambda_b R(b, d) \tag{1}$$

for some λ such that $\lambda_b > 0$, $\forall b$, and $\sum_b \lambda_b = 1$ is an admissible design. Further, when $\lambda_b = 1$, a design which minimizes (1) is admissible only if $R(b, d)$ has a unique minimum. Finally, there are no admissible designs other than these.

The function in (1) is the form used to obtain "compromise designs", and has been used extensively in an ad hoc manner (see references in Section 4). The admissibility framework, however, provides a more appropriate description of these compromise designs. The application of (1) to develop, in a systematic manner, defensible multiple objective designs is the major focus of this paper.

The following is a simple illustration of the method.

Example 1. Single-stage cluster sampling (with replacement): choice of selection probabilities.

Suppose one variable is to be measured for K subpopulations and that n out of N clusters are sampled. Also, once a cluster is sampled, every unit within the cluster is enumerated. Lastly, suppose a variable probability sample of clusters with replacement is taken. Then the expansion estimator of X_b, the population total, is

$$\hat{X}_b = \sum_{i=1}^{N} (\delta_i X_{bi} / n p_i)$$

where δ_i is a random variable denoting the number of times cluster i is selected in the sample, p_i is the probability of selecting cluster i in each draw, and X_{bi} is the population parameter for domain b, cluster i.
Assuming relative, quadratic loss,

$$R(b, d) = V(\hat{X}_b \mid d) / X_b^2 = \left\{ \sum_i ((X_{bi}/X_b)^2 / p_i) - 1 \right\} \Big/ n. \tag{2}$$

With a fixed total cost C and a cost of c_1 per primary unit, $n = C/c_1$.

The admissible values of the p_i, p_i^*, can be obtained by minimizing (1) where $R(b, d)$ is defined in (2).

For fixed $\boldsymbol{\lambda}$

$$p_i^* = \left(\sum_b \lambda_b (X_{bi}/X_b)^2 \right)^{1/2} \Bigg/ \sum_i \left(\sum_b \lambda_b (X_{bi}/X_b)^2 \right)^{1/2}. \tag{3}$$

Using the result stated below (1), these designs are admissible and they constitute all possible admissible designs.

If the measurement X_{bi} is proportional to the number of elements in the domain, M_{bi},

$$p_i^* = \left(\sum_b \lambda_b (M_{bi}/M_b)^2 \right)^{1/2} \Bigg/ \sum_i \left(\sum_b \lambda_b (M_{bi}/M_b)^2 \right)^{1/2}. \tag{4}$$

Suppose an estimate is needed only for domain t. Setting $\lambda_t = 1$, $p_i^* \propto M_{ti}$, which is the usual single purpose PPS design.

This example shows that when N is much larger than K, the set of admissible designs is much smaller than the set of candidate designs. While the set of candidate designs has dimension $N - 1$, the set of all admissible designs is a $(K - 1)$-dimensional subset of this larger set. Although a single optimal design cannot be obtained using the admissibility criterion alone, a (large) group of suboptimal designs have been ruled out.

In terms of p_i, $r(\boldsymbol{\lambda}, d)$ takes the same general form as each of the single purpose risks, $R(b, d)$ (see (1) for the definition of $r(\boldsymbol{\lambda}, d)$). In this case, knowing that a closed-form solution for the optimal design exists for each of the single purpose risks is an assurance that closed-form solutions exist for the admissible designs. In general, closed-form solutions can be obtained for other, more complex designs in which optimal, single purpose designs have a closed-form solution and the corresponding $r(\boldsymbol{\lambda}, d)$ has the same form as the single purpose risks (e.g., sample allocation for a stratified, two-stage design).

Example 2. Construction of strata for a sample of primary sampling units (PSUs).

The typical problem of multiple-objective stratification of PSUs is reported in Dahmstrom and Hagnell (1978) and in Kostanich et al. (1981). A sample is obtained by first selecting one PSU per stratum with probability proportional to a prespecified measure of size, then subsampling units within the sample PSU and finally recording a multivariate observation for each sampled second-stage unit. An estimate of the population total of each of the variables is desired. Suppose, further, that a fixed number of PSUs will be sampled so that a fixed number of strata L need to be formed. The problem in this case is to stratify the PSUs to produce precise estimates for each of the population totals. To form the strata these authors use only the between PSU component of variance (i.e., they ignore the within PSU variance).

Suppose there are $b = 1, \ldots, K$ different variables to be measured. For a specific stratification of all PSUs (implicitly stated as d), the between PSU variance corresponding to variable b is

$$R(b,d) = \sum_{h=1}^{L} \sum_{j=1}^{N_h} Z_{hj}(Y_{bhj}/Z_{hj} - Y_{bh})^2, \tag{6}$$

where, for stratification d, Y_{bhj} is the total for variable b, in PSU_j, of stratum h, N_h is the total number of PSUs in stratum h, $Y_{bh} = \sum_{j=1}^{N_h} Y_{bhj}$, A_{hj} is the measure of size for PSU (h,j) and $Z_{hj} = A_{hj}/\sum_{j=1}^{N_h} A_{hj}$ is the selection probability for PSU (h,j).

After, a reparametrization of (6), Dahmstrom and Hagnell (1978) provide an efficient search algorithm for finding the local minimum of $\sum_b R(b,d)$, using estimates for each Y_{bhj}. To achieve results closer to the global optimum, they suggest repeating the algorithm using different starting stratifications. Kostanich et al. (1981) also provide an efficient search algorithm which can find optimal or near optimal stratifications based on $\sum_b R(b,d)$ and other criteria. Clearly, the methods provided by these two papers can be used to produce admissible designs.

These examples illustrate how admissible designs can often be easily obtained for multiple objective surveys. This methodology may also be applied in choosing a sample design when experts disagree on the specifications of the finite population quantities. Using the admissibility criteria, and treating each expert's finite population specification as a different objective, an admissible set of designs can be obtained and their performance can be assessed. In addition, the robustness of a design can be assessed in a similar way by purposively entertaining different possible values of finite population parameters.

To illustrate the type of conclusion that can be drawn when using this methodology, Example 1 is investigated using 1980 population census data for Texas counties.

The sampling frame consists of all counties in Texas with a total population less than 50 000. It is desired to have estimates for the population covered by these counties as well as for the following subpopulations: Blacks, Hispanics, Asians and all persons aged 65 and over.

For this illustration, it is assumed that $X_{bi} \propto M_{bi}$, where M_{bi} is the total number of persons in subdomain b, county i. The values of the M_{bi} were obtained from the 1980 Census (US Census Bureau, 1981). If $M_{bi} = 0$, M_{bi} was taken to be 1. Using relative squared error loss, the admissible values of the p_i can be obtained from

$$p_i = \left\{ \sum_b \lambda_b (M_{bi}/M_{b.})^2 \right\}^{1/2} \Bigg/ \sum_i \left\{ \sum_b \lambda_b (M_{bi}/M_{b.})^2 \right\}^{1/2}$$

where $M_{b.} = \sum_i M_{bi}$.

In each part of Fig. 1, there is a separate curve for each of the five objectives (total, Black, Hispanic, Asian, aged). The vertical axis defines the CV of an estimate while the horizontal axis defines the value of λ. In Fig. 1(a), $\lambda' = (\lambda, 1-\lambda, 0, \ldots, 0)$ so that in (4), $\lambda_1 = \lambda$, $\lambda_2 = (1-\lambda)$ and $\lambda_3 = \lambda_4 = \lambda_5 = 0$. A point on the "Hispanic" curve gives the CV of the estimate for the Hispanic subpopulation corresponding to the value of λ on the horizontal axis. Although all 10 figures (corresponding to the distinct pairs of

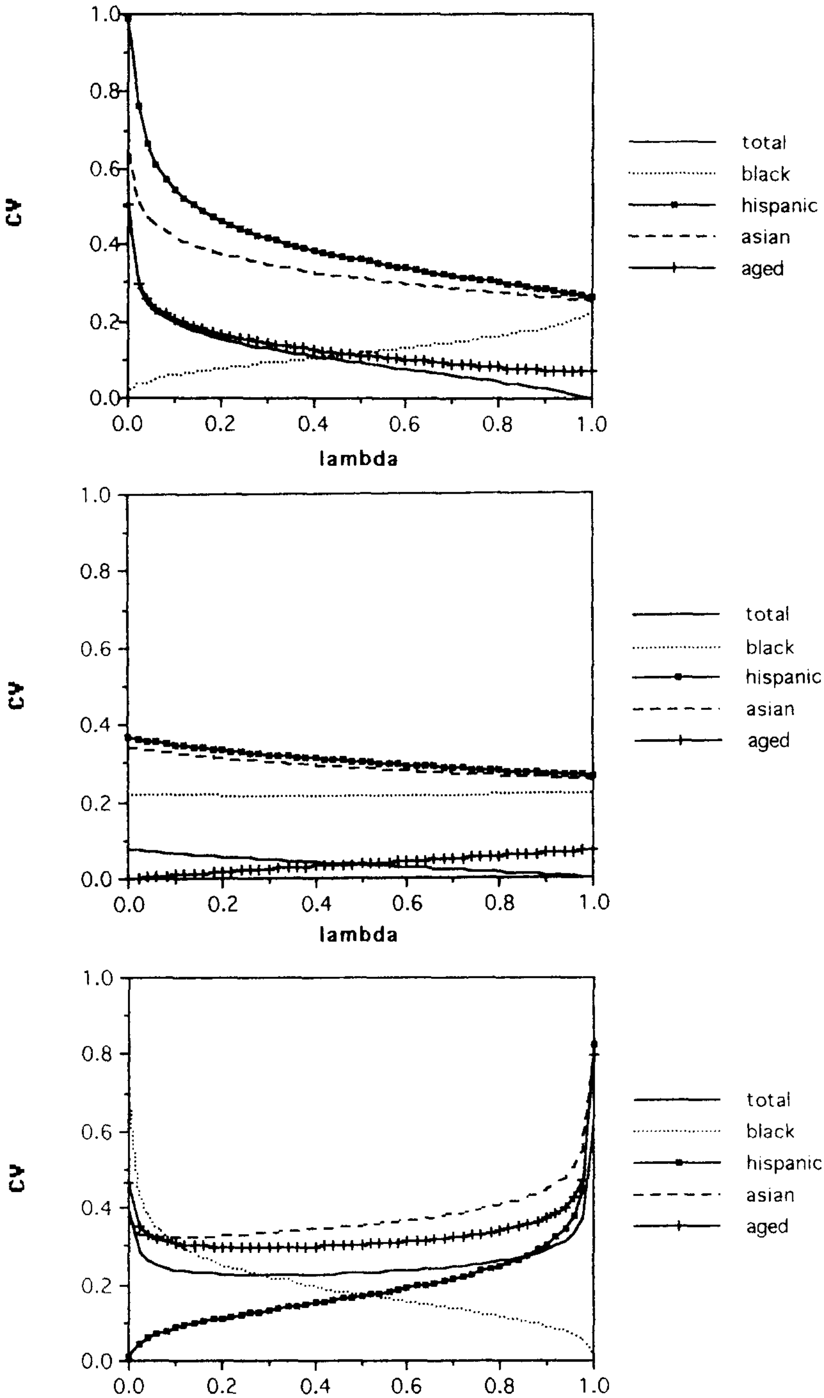

Fig. 1. Admissible designs based on two objectives a) total vs black tradeoff b) total vs aged tradeoff c) black vs hispanic tradeoff.

objectives) do not cover the entire space of admissible designs, they can be used to determine which pairs of objectives are compatible.

For example, suppose estimates corresponding to particular objectives are only of value if their CV is at most 0.05. Since Fig. 1(a) shows all admissible designs based on the pair of objectives, total population and Blacks, no design exists which will produce the desired CVs for both the total population and for Blacks. Similarly, Fig. 1(c) shows that there is no design which can produce estimates for both Blacks and Hispanics with the desired precision. However, in Fig. 1(b), it is seen that a design will produce precise estimates for both total population and the aged.

Suppose, instead, that a CV of 0.2 is an acceptable level of precision for estimates. Inspection of Fig. 1(a), shows that a design exists which will provide precise estimates for total population, the aged and Blacks. In Fig. 1(c), it can be seen that a design is available to estimate precisely both Blacks and Hispanics. The results presented in Figs. 1(a)–(c) cannot, however, rule out or confirm whether a design exists which will provide precise estimates for all five objectives. To investigate the existence of this design, one needs to search the admissible design space over all λ's. The next section includes an example of such a search.

3. Choosing a single design

The preceding section has shown how to select a class of designs, each member of which could be used as a final design. This section addresses the problem of choosing a final design from the admissible class.

There have been suggestions in the literature about how to pick a design when there are multiple objectives. For example, Yates (1960, Section 10.4) suggests weighting each component risk $R(b,d)$ by a monetary cost. This amounts to choosing one $\boldsymbol{\lambda}$, where λ_b is equal to monetary cost/unit of risk for objective b. A single design which minimizes the total monetary loss over all objectives is then selected.

Chatterjee (1967) defines the regret

$$P(b,d) = \left(R(b,d) - \min_d R(b,d) \right) \Big/ \min_d R(b,d)$$

and suggests minimizing $\sum_b P(b,d)$ for d.

This is equivalent to determining d in (1) with $\lambda_b = 1/\min_d R(b,d)$.

In stratified sampling, the power allocation (Bankier, 1988) is a method of allocating the sample of strata when separate estimates are desired for the strata, as well as for the entire population. Using this method, each stratum's component of the variance of the estimated population total receives a weight. The sample is then allocated by minimizing this weighted function. A range of weights is chosen so that one extreme corresponds to Neyman allocation and the other extreme corresponds to an allocation yielding equal CVs for the estimated strata means. These extreme weights and all of the intermediate weights correspond to a subset of the admissible solutions in (1). Folks and Antle (1965) developed admissible designs for allocation of sample sizes to strata but, apparently, never exploited the technique as a method for designing multiple objective surveys.

As discussed in Section 2, in a multivariate, stratified, two-stage, probability proportional to size sample setting, Dahmstrom and Hagnell (1978) and Kostenich et al. (1981) provide efficient search algorithms to determine a stratification that optimizes, or nearly optimizes, (1). In addition, Kostenich et al. (1981) use their algorithm to find stratifications that optimize, or nearly optimize, criteria other than a weighted average of risks. Stratifications based on these other criteria should be checked as to whether they are admissible or not.

To determine strata sample sizes, Cochran (1977, Section 5A.3) suggests optimizing each objective separately and then averaging the resulting strate sample sizes to obtain a compromise design. In two-stage sampling, Folsom et al. (1987) average measures of size to define first-stage selection probabilities. In general, procedures like these may not yield an admissible design. As an illustration, the optimal choice of selection probabilities for a single purpose b (Example 1) is

$$p_{bi} = M_{bi} \Big/ \sum_i M_{bi}.$$

A simple average yields a compromise design

$$p_i^+ = \frac{1}{K}\sum_{b=1}^{K}\left(M_{bi} \Big/ \sum_i M_{bi}\right).$$

This design is admissible only if there exists a λ, in the K-dimensional simplex such that $p_i^+ = p_i^*$; that is, when

$$\frac{1}{K}\sum_{b=1}^{K}\left(M_{bi} \Big/ \sum_i M_{bi}\right) = \left(\sum_b \lambda_b M_{bi}^2\right)^{1/2} \Big/ \sum_i \left(\sum_b \lambda_b M_{bi}^2\right)^{1/2} \quad \text{for } i = 1,\ldots,N-1. \tag{7}$$

Since (7) requires solving a system of $N-1$ equations in $K-1$ unknowns, there will be cases when no solution exists and the compromise design, based on an average of selection probabilities, will be inadmissible.

The only sure way to pick a final satisfactory design is to evaluate $L(b, d_\lambda)$ for each b and all λ and then determine which λ produces the most acceptable trade-off among the resulting risks. If the individual loss functions are comparable to each other, as with relative squared error loss, more automatic procedures are available. One way is to search the admissible design space for the minimax design (i.e., the design which minimizes the maximum risk function over all objectives).

Five such searches were conducted for Example 1. First, the λ which minimized the maximum CV over all five objectives was determined. Next, the minimax design for each subset of four objectives was obtained and the design corresponding to the minimum of these minimax designs was selected. Continuing in this manner, a

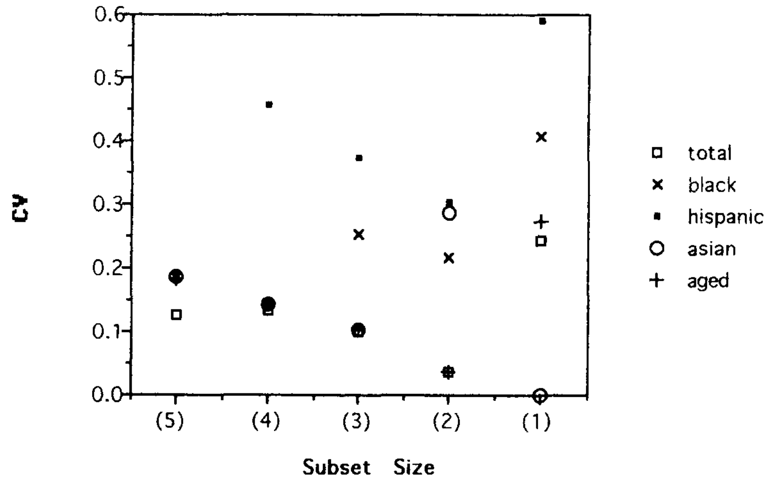

Fig. 2. Minimax and subset minimax designs.

Table 1
Values of λ for minimax and subset minimax designs

Objective	Design				
	minimax	mini-(4)	mini-(3)	mini-(2)	minimin
Total	0.0584	0.096	0.0524	0.4832	0
Black	0.3889	0.4552	0	0	0
Hispanic	0.3448	0	0	0	0
Asian	0.2072	0.3560	0.4600	0	1
Aged	0.0008	0.0928	0.4876	0.5168	0

minimum, minimax design was determined for subsets of size three, two and finally one. Fig. 2 summarizes these results. The abscissa denotes the minimum, minimax design for subsets of size five, four, three, two and one. For each of these five designs, the corresponding CVs for each of the five objectives are plotted on the ordinate.

The minimax design in the leftmost side of Fig. 2 shows that there is a design where each of the five objectives has a CV less than 0.2. The next set of CVs in the figure shows that a design which will produce a CV of less than 0.15 can be obtained for four of the five objectives (dropping Hispanics). On the far right of the figure, the design which minimizes the minimum CV would produce estimates for Asians with a CV of the order of 10^{-6}. A table of this type can be used by survey planners to assess further tradeoffs in providing estimates for all or a subset of objectives. By first identifying the set of admissible designs, the search is greatly simplified because of the smaller design space and because the search can be conveniently indexed by λ in a simplex.

Suppose that in this example a CV of 0.2 was unacceptable. Fig. 2 shows that sacrificing a good design for Hispanic estimates will produce the greatest reduction in CV for the remaining four objectives.

Table 1 shows the λ's for each of the five designs represented in Fig. 2. Contrary to what might be expected, the minimax design does not yield a design based on equal weights.

5. Summary

The admissibility criterion has been shown to be a useful way to reduce the class of candidate designs for a multiple objective survey, given a fixed cost. In this context, the admissibility criterion is defined over the range of survey objectives, unlike the usual situation where admissibility is defined over the parameter space. Manageable methods for choosing a single design from a class of admissible designs were presented. One way is to pick a minimax design or pick a design which is minimax for a more realistic subset of designs. Another way is to graph the tradeoffs between competing designs. Examples were presented which illustrate the use of both the admissibility and minimax criteria. Other methods for obtaining compromise designs were reviewed and discussed, many of these methods yielding designs which are in an admissible class of designs. In particular, the practice of optimizing a weighted average of separate risk functions was shown to yield an admissible design.

Acknowledgements

The author would like to thank the associate editor and the referees for providing valuable comments that improved the presentation of the paper.

References

Bankier, M.D. (1988). Power allocations: determining sample sizes for subnational areas. *Amer. Statist.* **42**, 174–177.

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.

Census of Population and Housing (1980). Summary tape file 1A (Texas) [machine-readable data file]/prepared by the Bureau of the Census. The Bureau [producer and distributor], Washington, 1981.

Chatterjee, S. (1967). A note on optimum stratification. *Skand. Akt.* **50**, 40–44.

Cochran, W.G. (1977). *Sampling Techniques*. Wiley, New York.

Dahmstrom, P. and M. Hagnell (1978). Multivariate stratification of primary sampling units in a general purpose survey. *Scand. J. Statist.* **5**, 53–56.

Folks, J.L. and C.E. Antle (1965). Optimum allocation of sampling units when there are *R* responses of interest. *J. Amer. Statist. Assoc.* **60**, 225–233.

Folsom, R.E., F.J. Potter and S.R. Williams (1987). Notes on a composite size measure for self-weighting samples in multiple domains. *Proc. on Survey Research Methods*, American Statistical Association, 792–796.

Huddleston, H.F., P.L. Claypool and R.R. Hocking (1970). Optimal sample allocation to strata using convex programming. *Appl. Statist.* **19**, 173–278.

Kokan, A.R. (1963). Optimum allocation in multivariate surveys. *J. Roy. Statist. Soc. Ser. A* **126**, 557–565.

Kostanich, D., D. Judkins, R. Singh and M. Schautz (1981). Modification of Friedman–Rubin's clustering algorithm for use in stratified PPS sampling. *Proc. Survey Research Methods*, American Statistical Association, 285–290.

National Center for Health Statistics, C.A. Schoenborn and M. Marano (1988). Current estimates from the National Health Interview Survey: United States, 1987. *Vital and Health Statistics. Series* 10, *No.* 166. DHHS Pub. No. (PHS) 88–1594, Public Health Service, U.S. Government Printing Office, Washington.

Yates, F. (1960). *Sampling Methods for Censuses and Surveys.* Charles W. Griffin, London.