Optimal scheduling of contact attempts in mixed-mode surveys

Strategy allocation for adaptive survey designs

Calinescu, Melania (1st author)
VU University Amsterdam, Stochastics
De Boelelaan 1081
Amsterdam 1081HV, The Netherlands
E-mail: melania.calinescu@vu.nl

Bhulai, Sandjai (2nd author)
VU University Amsterdam, Stochastics
De Boelelaan 1081
Amsterdam 1081HV, The Netherlands
E-mail: s.bhulai@vu.nl

1 Introduction

Recent survey literature shows an increasing interest in data collection methods that tailor strategies and resources to subgroups in the target population. The reason behind it is that current survey designs use uniform strategies that do not differentiate efforts, although it has been shown that effectiveness of different survey design features and causes for nonresponse vary greatly over persons and households. Apart from the intrinsic need for such flexible designs, the change of focus is also driven by technological advances and potential savings in survey budgets. Survey management and monitoring systems can be upgraded in order to allow for segmentation of samples over different strategies. Furthermore, web surveys are substantially cheaper than interviewer-administered modes, which raises the question of whether surveys can be much cheaper without a strong decrease in quality. This paper presents a case study in which we differentiate efforts in order to optimize quality given constraints on costs.

There are various factors that influence the decision to accept or reject the request to participate in sample surveys. Such factors include the social characteristics of the sample unit, the interviewer's behavior, and the attributes of the survey design, e.g., the interview mode, the schedule of the contact attempts, and the language of the questionnaire. Traditional survey designs do not use the information that results from studying these factors. Predicting the probability that a sample unit would respond given a list of factors, quantifying the factors' effects on human behavior in terms of costs and data quality can be of great help in designing a high-quality but cost-effective survey.

While the research in the field is still in its infancy, there have been some design methods investigated or implemented. The first attempts belong to [1], where the authors describe a two-phase sampling framework for nonresponse. This approach has been termed as *responsive survey design*. More advanced methods in the same context of *responsive design* are later presented in [2] and [4], where the main idea is to identify a set of design attributes that potentially influence the survey costs and the errors in the survey estimates and to monitor their influence on costs and errors. This information helps in subsequent phases to alter the design features such that a desired balance between costs and errors is achieved. Another example is given in [3], where the authors investigate the influence on cooperation and costs of the number of call attempts.

The historical data considered in the aforementioned studies do not depend on person or household characteristics, the use of which could bring improvements in the response rate predictions. Such characteristics are available from administrative registers, financial records, etc. When such information is employed to adjust the design features for a given set of characteristics (i.e., different design features can be applied to different sample units) the resulting survey design is termed *adaptive survey design* and it has been generally introduced in [8] and [6]. The method originates from the field of clinical trials, where treatments are group-specifically set before the start of the trial and change during the trial according to the responses of patients. Therefore, the great advantage of this method lies in its flexibility. It is tailored for groups of sample units, it can be defined before the survey starts and also updated during the data collection based on information regarding the characteristics of respondents and nonrespondents.

In the present paper, adaptive survey designs are analyzed from the perspective of resource allocation problems, which constitutes the novelty of this research. Given a budget, a set of household characteristics, and a list of factors that influence the survey costs and quality, we develop a model that computes the allocation of survey resources such that quality is maximized while costs meet the budget constraint. Extracting detailed information from historical data and building a survey design that is both cost-efficient and of high-quality are the main contributions the current paper brings to practice and the research in the field.

The remainder of the paper is structured as follows. Section 2 gives an overview of the concepts in adaptive survey designs and Section 3 discusses the mathematical model and an algorithm to derive optimal adaptive survey design policies. Numerical examples of these situations are given in Section 3.4. Section 4 concludes the results of the paper and gives directions for future research.

2 An adaptive survey design model

2.1 Notation

Let the population consist of units $k = 1, 2, \ldots, N$. The population of interest may consist of all units in a population but also of all respondents to a previous wave in a panel study. To each unit a strategy s will be assigned from the set of candidate strategies $\mathcal{S} = \{\emptyset, s_1, \ldots, s_M\}$. Strategy $\emptyset$ implies that no action is undertaken, i.e., the population unit is not sampled. In this paper, however, the sampling design is not part of the strategy allocation, the sample is given and fixed. However, one may include the decision to sample a unit explicitly in the overall allocation of resources.

In general, a strategy s is a specified set of design features, e.g., s_1 = *(advance letter, web questionnaire, one reminder)*. A strategy may involve a sequence of treatments where treatments are only followed when all previous treatments failed, e.g., s_2 = *(advance letter, web questionnaire, one reminder, CATI questionnaire, maximum of six call attempts)*. It is assumed in this paper that the set of strategies $\mathcal{S}$ is known and fixed when the strategy allocation starts. The set of strategies may be identified based on historical survey data, experience and pilot studies.

The data collection can vary for different sample units. For example, follow-ups for refusal conversion can increase the fieldwork for specific units. Most common survey modes are *paper*, *web*, *telephone* and *face-to-face*. Each mode carries different levels of costs and quality which suggests that different modes should be allocated to different sample units (i.e., a mixed-mode design). The contact protocol has a significant impact on the quality of the survey results. Different number of contact attempts or a specific timing for these attempts can influence the willingness to participate in the survey.

Sample units can be clustered into homogeneous groups based on characteristics such as age, gender, ethnicity, information that are available for all units from external sources of data. Let $\mathcal{G} = \{1, \ldots, G\}$ be the set of homogeneous groups with size N_g for group $g \in \mathcal{G}$ in the survey sample. The proportion or importance of group g in the sample is then given by w_g (usually $w_g = N_g/N$). A second set of characteristics, called *paradata*, may be available for unit k gathered during data collection (e.g., the interviewer assessment of the propensity to respond). If information from paradata is considered for clustering the sample units then the adaptive design is termed *dynamic*. In the current paper we consider only *static* adaptive designs, i.e., clustering is based only on information available for all sample units from external sources. However, the approach can be extended to include paradata.

Let $\tau_g(s)$ be the expected response propensity for units in group g that are assigned strategy s. Such propensities are estimated from historical data, e.g., previous versions of the same survey, surveys with similar topics and designs. It is imperative that response propensities are estimated from randomized contrasts, i.e., historic data must carry randomized assignments to different design features. For example, from a face-to-face survey with a maximum of three visits it cannot be inferred what the response propensities would be if the maximum is raised to ten visits. Similarly, in telephone surveys with evening calls only, it cannot be deduced what the contact propensities would be for daytime shifts. Randomization over design features is, however, not sufficient. Historical data also need randomization over groups in $\mathcal{G}$.

There are various models in the literature that estimate response propensities while taking into consideration survey design features, such as nested or sequential regression models, multilevel models. However, uncertainty is still present in these estimates due to sampling variation and design mismatch between similar surveys. Therefore, sensitivity analysis should always be used in order to obtain insight into the variation of quality and costs caused by randomness in the input probabilities.

Let $\rho(s|g)$ be the allocation probability of strategy s to a sample unit from group g. This implies that

$$0 \leq \rho(s|g) \leq 1, \ \forall s \in \mathcal{S}, g \in \mathcal{G},$$
$$\sum_{s \in \mathcal{S}} \rho(s|g) = 1, \ \forall g \in \mathcal{G}.$$

Allowing for allocation probabilities between 0 and 1 increases the flexibility in meeting quality levels or cost constraints.

To completely define the problem at hand we need to specify the quality and the cost functions. It is important to keep in mind that other survey errors may sometimes play a dominant role and that design features that have a high risk of survey errors other than nonresponse must be avoided. Such design features are survey modes with low coverage of some groups or a higher risk of response bias, increased item nonresponse in proxy reporting or increased response bias in follow-ups.

In general two types of quality functions can be distinguished, i.e., quality functions based on characteristics from external data and paradata only, and quality functions that employ additionally the answers to the survey target variables. The latter asks for a model-based approach since answers from nonrespondents are missing. We focus on the first category of functions. Two well-known examples are the response rate

$$(1) \quad \bar{\tau} = \sum_{g \in \mathcal{G}} \sum_{s \in \mathcal{S}} w_g \, \tau_g(s) \, \rho(s|g),$$

and the representativeness indicator

$$(2) \quad R = 1 - 2\sqrt{\sum_{g \in \mathcal{G}} (\bar{\tau}_g - \bar{\tau})^2},$$

where $\bar{\tau}_g = \sum_{s \in \mathcal{S}} w_g \, \tau_g(s) \, \rho(s|g)$ and $\bar{\tau}$ is as given by (1).

Denote by $c_g(s)$ the expected costs from assigning strategy s to group g. Note that startup costs (e.g., designing questionnaires) are not taken into consideration. We assume here that such a cost function can be constructed based on historical data, i.e., costs can be estimated for a given survey, set of strategies and sample groups. The total costs for handling a sample are computed as

$$(3) \quad \sum_{g \in \mathcal{G}} \sum_{s \in \mathcal{S}} w_g \, c_g(s) \, \rho(s|g).$$

For a more detailed description of adaptive survey designs we refer the reader to [6] and [8].

2.2 Numerical example

We illustrate the static setting of adaptive designs by means of an example. Suppose the strategy set is given by $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$, with

s_1 = (web administered, 1 reminder);

s_2 = (web administered, no reminders, CAPI administered, maximum 3 visits);

s_3 = (CAPI administered, maximum 6 visits);

s_4 = the empty strategy.

Strategies s_1 and s_3 lead to a single-mode design while s_2 describes a mixed-mode survey where CAPI is administered to the nonrespondents from the initial wave of web.

We consider a sample of size $N = 2,000$ clustered in two groups given an age criterion: g_1 groups sample units with age below or equal to 35 and g_2 units with age above 35. The proportion of groups in the sample is given by $q(x) = (0.5, 0.5)$. With (1) as our objective, the task at hand is to assign survey modes to groups such that the average response is maximized.

Costs are induced every time one of the following actions is taken

web questionnaire	:	€ 5;
web reminder	:	€ 2;
one visit	:	€ 15;
one interview	:	€ 20.

In order to investigate the trade off between quality and costs we set up a budget range between 0 and € 50,000 and require that costs do not overrun the budget. Moreover, we assume that there is enough interviewer capacity to carry out the survey.

The response propensities $\tau_g(s)$ are estimated from historical data (see Table 1). On the basis of these estimates we compute the expected costs $c_g(s)$ given in Table 1. Given this input we can

Table 1: Response propensities and expected costs

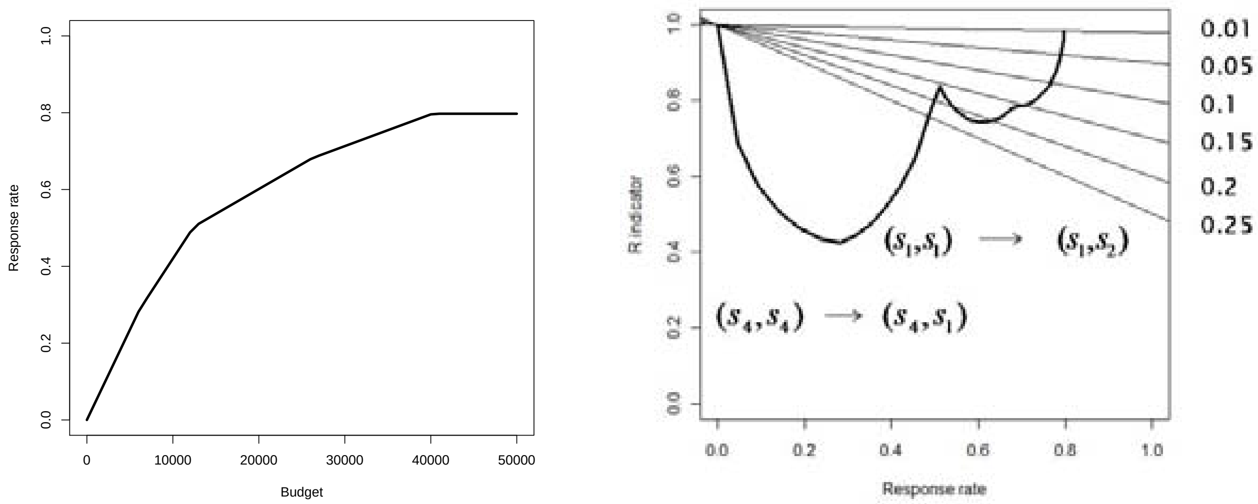
	g_1				g_2			
Strategy	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4
Response	0.432	0.789	0.684	0	0.574	0.806	0.663	0
Costs	6,320	20,094	20,535	0	6,120	20,103	23,132	0

conclude that s_2 is preferred if enough budget is available. Moreover, assigning s_3 to either group is not optimal since it brings a lower response rate than s_2 with greater costs.

The strategy allocation probabilities $\big(\rho(s|g)\big)_{s\in\mathcal{S},g\in\mathcal{G}}$ form the set of decision variables. The optimal solution with the corresponding objective value $\bar{\tau}$ is given in Table 2 for the indicated budget levels. Using the definition given by (2) we compute the R-indicator for each budget level. A budget larger than € 40,200 does not lead to further improvements in the response rate. As expected, when resources are scarce the empty strategy is employed. With the gradual increase of the budget the strategy allocation becomes a mixture of s_1 and s_4, s_1 and s_2 and ultimately with probability 1 s_2 is assigned to both groups when enough budget is available. Figure 1(a) displays the evolution of the average response for various budget levels. A response-representativeness function is proposed in [7] to reflect acceptable levels for R-indicator and response rates. The function $R(\bar{\tau})$ has an acceptable level if $R(\bar{\tau})$ exceeds a given maximal bias threshold. In Figure 1(b) the lines represent the maximal

Table 2: Optimal solution and corresponding response rate and R-indicator

	Budget						
	0	5,000	10,000	20,000	30,000	40,000	40,200
$\rho(s_1\|g_1)$	0	0	0.614	0.451	0	0	0
$\rho(s_2\|g_1)$	0	0	0	0.549	1	1	1
$\rho(s_3\|g_1)$	0	0	0	0	0	0	0
$\rho(s_4\|g_1)$	1	1	0.386	0	0	0	0
$\rho(s_1\|g_2)$	0	0.817	1	1	0.729	0.014	0
$\rho(s_2\|g_2)$	0	0	0	0	0.27	0.986	1
$\rho(s_3\|g_2)$	0	0	0	0	0	0	0
$\rho(s_4\|g_2)$	1	0.183	0	0	0	0	0
$\bar{\tau}$	0	0.235	0.420	0.601	0.713	0.796	0.797
R	1	0.435	0.571	0.743	0.790	0.959	0.983



(a) Response rate for various budget levels (b) Response-representativeness function

Figure 1: Objective value evolution for various budget levels

bias thresholds with the indicated values. Once $R(\bar{\tau})$ exceeds the line, the corresponding maximal bias threshold is met. The drops in the $R(\bar{\tau})$ function are caused by the different mixtures of allocated strategies.

3 Optimal scheduling of contact attempts

In the following the objective is to maximize the response rate given by (1) while tailoring two design features, namely *survey modes* and *contact protocol*, in particular the *maximum number of contact attempts per survey mode*. Strategies in $\mathcal{S}$ have various sequences of survey modes (face-to-face, phone, web, paper) and various values for the number of contact attempts (in case of interviewer-assisted modes). We know that for interviewer-assisted modes successful participation in the survey depends on first establishing contact, and then cooperation by answering the questionnaire. Therefore, we view the response propensity $\tau_g(s)$ through its components, i.e., contact probability p and cooperation probability r. To maintain a general framework, we assume that for self-administered modes the contact probability is equal to 1.

3.1 Notation

Consider the fieldwork divided into time slots (i.e., days in a month, shifts in a day) at which units in a group can be approached for a survey. Denote by $\mathcal{T} = \{1, \ldots, T\}$ the set of time units. The set of different survey modes is denoted by $\mathcal{M} = \{1, \ldots, M\}$. At each time slot $t \in \mathcal{T}$ one can decide to approach units in group $g \in \mathcal{G}$ for a survey using mode $m \in \mathcal{M}$. From historical data group-dependent contact probabilities $p_g(t, m)$ and response probabilities $r_g(t, m)$ can be estimated, which we consider as given quantities in our problem. Note that from historical data it can also be observed that certain time slots (e.g., morning, evening) have an influence on the availability of the unit and the willingness to respond. Therefore, to employ most of the available information, the contact and response probabilities are modeled at the level of time slots for each group and per mode.

When a successful contact is established and the unit responds positively, the survey ends with success; this happens with probability $p_g(t, m)r_g(t, m)$. However, if the unit responds negatively after successful contact, the unit is not considered for a future survey approach; this happens with probability $p_g(t, m)\big(1-r_g(t, m)\big)$. Only in the case that the unit is not contacted successfully, the unit can be considered for a future survey approach; this happens with probability $1 - p_g(t, m)$. Thus, if the unit is approached again at time t' using mode m', then the probability of a successful approach is $\big(1-p_g(t, m)\big)p_g(t', m')r_g(t', m')$, and the probability of a contact failure is $\big(1-p_g(t, m)\big)\big(1-p_g(t', m')\big)$.

Denote by $x_g(t, m)$ a binary 0-1 decision variable that denotes if units in group g are approached for a survey at time t using mode m. Then, the probability that a contact fails at time t', denoted by $f_g(t')$, is given by

$$f_g(t') = \prod_{t=1}^{t'} \prod_{m \in \mathcal{M}} \left[x_g(t, m)\big(1 - p_g(t, m)\big) + \big(1 - x_g(t, m)\big)\right].$$

This is a highly non-linear expression in the decision variables, which can be recursively computed by

$$(4) \qquad f_g(t') = \prod_{m \in \mathcal{M}} \left[x_g(t', m)\big(1 - p_g(t', m)\big) + \big(1 - x_g(t', m)\big)\right] f_g(t' - 1),$$

using the fact that $f_g(0) = 1$.

3.2 Model

Note that values of $\big(x_g(t, m)\big)_{t\in\mathcal{T}, m\in\mathcal{M}}$ form a set of matrices for each $g \in \mathcal{G}$. An element of this set identifies a strategy. Using definition (4) and the fact that each group g is assigned one single strategy the group response rate $\tau_g(s)$ can then be computed by

$$\sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} f_g(t - 1)\, x_g(t, m)\, p_g(t, m)\, r_g(t, m).$$

Hence, the objective of the decision maker becomes to maximize

$$(5) \qquad \sum_{g\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} w_g\, f_g(t - 1)\, x_g(t, m)\, p_g(t, m)\, r_g(t, m),$$

by setting the decision variables $x_g(t, m)$ optimally.

The decision variables are subject to constraints, though, due to scarcity in resources. Note that a maximum number of attempts per sample unit is required in order to manage the costs and interviewer capacity. If a budgetary constraint and a capacity limitation are added to the model then such a number becomes obsolete.

Let us define these two constraints. Both budget and capacity usage depend mainly on the interview mode and on the outcome of each approach. Denote by $b^s(m)$ the costs that are incurred by using mode m with a successful outcome. For the costs that are incurred by mode m that results in a failure, we distinguish two types of costs: $b^{f_c}(m)$ when the failure occurs due to failure of contact, and $b^{f_r}(m)$ when the failure occurs due to failure to respond. Let B be the total budget that is available for the survey. An approach at time t using mode m bears the following costs

$$p_g(t,m)\big[r_g(t,m)b^s(m) + \big(1 - r_g(t,m)\big)b^{f_r}(m)\big] + \big(1 - p_g(t,m)\big)b^{f_c}(m).$$

In general, the costs $b_g(t,m)$ at time t using mode m depend on the contact failures before time t. These costs can be written as follows

$$b_g(t,m) = x_g(t,m)f_g(t-1)\Big[p_g(t,m)\big[r_g(t,m)b^s(m) + \big(1 - r_g(t,m)\big)b^{f_r}(m)\big] + \big(1 - p_g(t,m)\big)b^{f_c}(m)\Big],$$

with $f_g(t)$ given by (4). Hence, using this definition, the budgetary constraint that needs to be added to our model is given by

$$\sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}} N_g b_g(t,m) \le B.$$

A capacity constraint can be addressed in a manner analogous to the constraint on the budget. Let C be the available capacity, measured by the number of interviewer hours available to survey the sample. Similar to the budgetary cost structure, the required capacity depends on the interview mode and on the outcome of each approach. Denote by $c^s(m)$, $c^{f_c}(m)$, and $c^{f_r}(m)$ the capacity utilized when the approach is successful, or has failed due to contact failure, or failed due to a nonrespondent, respectively. Following the same steps as above, the capacity constraint to be added to the model is given by

$$\sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}} N_g c_g(t,m) \le C,$$

with $c_g(t,m)$ defined as

$$c_g(t,m) = x_g(t,m)f_g(t-1)\Big[p_g(t,m)\big[r_g(t,m)c^s(m) + \big(1 - r_g(t,m)\big)c^{f_r}(m)\big] + \big(1 - p_g(t,m)\big)c^{f_c}(m)\Big].$$

Furthermore, at time t only one mode can be employed to approach a group, yielding the constraint $\sum_{m\in\mathcal{M}} x_g(t,m) \le 1$. By combining the objectives with all the constraints, we can draft our optimization problem as a binary programming problem in the following manner.

$$\begin{aligned}
\max\ & \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}} w_g\, f_g(t-1)\, x_g(t,m)\, p_g(t,m)\, r_g(t,m) \\
\text{s.t.}\ & \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}} N_g b_g(t,m) \le B, \\
& \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}} N_g c_g(t,m) \le C, \\
& \sum_{m\in M} x_g(t,m) \le 1,\ \forall g\in\mathcal{G},\ \forall t\in\mathcal{T}, \\
& f_g(t) = \prod_{m\in\mathcal{M}} [x_g(t,m)(1 - p_g(t,m)) + 1 - x_g(t,m)] f_g(t-1),\ \forall g\in\mathcal{G},\ \forall t\in\mathcal{T}, \\
& f_g(0) = 1,\ \forall g\in\mathcal{G}, \\
& x_g(t,m) \in \{0,1\},\ \forall g\in\mathcal{G},\ \forall t\in\mathcal{T},\ \forall m\in\mathcal{M},
\end{aligned} \tag{6}$$

with

$$b_g(t,m) = x_g(t,m) f_g(t-1) \Big[p_g(t,m) \big[r_g(t,m) b^s(m) + \big(1 - r_g(t,m)\big) b^{f_r}(m) \big] + \big(1 - p_g(t,m)\big) b^{f_c}(m) \Big], \ \forall g \in \mathcal{G}, \ \forall t \in \mathcal{T}, \ \forall m \in \mathcal{M},$$

$$c_g(t,m) = x_g(t,m) f_g(t-1) \Big[p_g(t,m) \big[r_g(t,m) c^s(m) + \big(1 - r_g(t,m)\big) c^{f_r}(m) \big] + \big(1 - p_g(t,m)\big) c^{f_c}(m) \Big], \ \forall g \in \mathcal{G}, \ \forall t \in \mathcal{T}, \ \forall m \in \mathcal{M},$$

Note that in this formulation, we have chosen to model the budgetary constraint and the capacity restriction as a global constraint over all the groups. However, it is quite easy to divide the budget B into budgets B_g for each group g, and then have a constraint per group. A similar remark holds for the capacity restriction as well.

3.3 Algorithm

The solution of problem (6) is, however, not trivial. The objective function is a non-convex non-linear function, and the constraints do not form a convex polytope either. As a consequence, our problem is non-tractable even for small-sized problems (e.g., 1 group and 4 time slots!). However, this issue is eliminated if (6) can be formulated as a Markov decision problem. In order to check this, note that at any time t, it is sufficient to know $f_g(t)$ instead of the complete configuration $x_g(t',m)$ for $t' \leq t$ for all g. Hence, given $f_g(T)$, the decision at time T is obvious when one also keeps track of the number of times that mode m has been used for each group g. Since the decision at time T is completely determined, one can then calculate the optimal decisions at time $T-1$, and continue working back towards the first time epoch. By keeping track of the time, the contact failure probability, and the utilization of the different modes, the problem becomes completely Markovian and can be cast as a Markov decision problem.

A Markov decision problem (see also [5]) can be solved by dynamic programming (or backward recursion) with guaranteed convergence to the optimal solution. Note that the algorithm only needs T iterations, and in each iteration only $2^{G \cdot M}$ actions need to be considered. Hence, for values of realistic size, the algorithm is computationally feasible.

The optimal solution describes per group the best balance between various survey features. Then $\rho_g(s)$ will be equal to 1 for strategy s that matches this balance. If no such strategy s exists then the final solution is given by a linear combination of closely related strategies (i.e., design features have values close to those indicated by the optimal solution) such that the optimal quality level is achieved and constraints are fulfilled. Then the allocations probabilities of the respective strategies are equal to the coefficients of the linear combination thus computed.

3.4 Numerical example

The previous sections dealt with the theoretical model to solve the problem of resource allocation within static adaptive survey designs. In this section, we illustrate our methodology by means of a numerical example. The focus is on the optimal sequence of various modes.

In this setting, we drop the budget and the capacity constraints and we look instead at the maximum number of attempts per mode. The example shows that the solution to this model is indeed optimal, although, counterintuitive. Consider a survey sample in which all units belong to the same group g. The set of available interview modes is $\mathcal{M} = \{\text{face-to-face, phone}\}$. The survey fieldwork is divided in $T = 6$ time slots. Table 3 gives the contact and response probabilities $p_g(t,m)$ and $r_g(t,m)$ as estimated from previous such surveys and the maximum number of attempts $\bar{k}_g(m)$ for mode $m \in \mathcal{M}$.

Table 3: Input data for group g

Mode	Probability	t_1	t_2	t_3	t_4	t_5	t_6	$\bar{k}_g(m)$
Face-to-face	$p_g(t,m)$	0.3	0.4	0.8	0.2	0.3	0.7	2
	$r_g(t,m)$	0.9	0.7	0.3	0.8	0.8	0.6	
Phone	$p_g(t,m)$	0.4	0.5	0.9	0.4	0.4	0.8	4
	$r_g(t,m)$	0.8	0.5	0.7	0.6	0.4	0.6	

Note that there is a clear preference for contact at time slots t_3 and t_6 for both interview modes. For response, on the other hand, the probabilities indicate more than 50% chance for positive response except for an attempt by face-to-face at t_3 and by phone at t_5. Therefore, it is not obvious what time slots should be chosen in order to maximize the total reward.

The strategy set is $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$, with

$$s_1 = (T = 6,\ \text{face-to-face, 2 attempts, phone, 4 attempts})$$
$$s_2 = (T = 6,\ \text{face-to-face, 3 attempts, phone, 4 attempts})$$
$$s_3 = (T = 6,\ \text{face-to-face, 2 attempts, phone, 3 attempts})$$
$$s_4 = (T = 6,\ \text{face-to-face, 3 attempts, phone, 3 attempts}).$$

Using the algorithm from Section 3, we obtain the solution depicted in Table 4.

Table 4: Optimal solution – original setting

Time slot	t_1	t_2	t_3	t_4	t_5	t_6	Response rate
Mode	F2F	F2F	Ph	Ph	0	Ph	0.753

Let us analyze this solution. It looks surprising that for the first time slot mode F2F is chosen and not Ph, although the immediate reward is higher for Ph. However, considering the formula given in (4) for the group average response, we see that the lower the contact probability for the first time slot, the higher the future reward. Also, the response probability $r_g(t_1, \text{F2F})$ is higher than $r_g(t_1, \text{Ph})$. The situation changes when $r_g(t_1, \text{F2F}) < r_g(t_1, \text{Ph})$. For example, take $r_g(t_1, \text{F2F}) = 0.7$. As expected, the new optimal solution (see Table 5) uses *phone* as first approach interview mode.

Table 5: Optimal solution – different response probability at t_1

Time slot	t_1	t_2	t_3	t_4	t_5	t_6	Response rate
Mode	Ph	F2F	Ph	Ph	F2F	Ph	0.736

The structure of the solution given in Table 4 is motivated by the choice of $\bar{k}_g(m)$. From t_3 onward *phone* is the only interview mode left available. Thus, the choice for time slots t_3, t_4, and t_6 is logical. However, taking action 0 at t_5 again looks counterintuitive. Since there are enough attempts left for mode Ph and there are no budget or capacity constraints, it feels natural to choose for an attempt to approach. The explanation lies in the value of the objective function that is higher in this case (0.753 compared to 0.752 if the unit is approached).

The optimal solution in Table 4 does not employ all attempts available for mode Ph. Therefore, we cannot obtain a different solution if we increase the number of attempts for this mode. Consequently, selecting strategies s_1 or s_2 is not necessary. On the other hand, if we increase the number of attempts to 3 for F2F, then the average response improves (see Table 6). The structure of the optimal solution does not change much from the original setting. The only difference appears at t_5 where this time there are enough attempts for mode F2F, and selecting this mode leads to a higher reward.

Table 6: Optimal solution – more attempts available

Time slot	t_1	t_2	t_3	t_4	t_5	t_6	Response rate
Mode	F2F	F2F	Ph	Ph	F2F	Ph	0.755

We conclude that given the set $\mathcal{S}$ strategy s_4 gives the maximum response rate, thus $\rho(s_4|g) = 1$.

4 Conclusions

For every survey that is planned, survey organizations are confronted with the decision over the necessary budget such that the resulting survey quality is above a pre-agreed level. In most cases, the budget proves insufficient due to an increased effort to convince sample units to respond. Temporary solutions can be found in replacing the expensive design with a cheaper one. That, however, has a negative influence on the survey quality. Moreover, the decreasing response prevents expensive designs from performing as well as it is expected. Therefore, a new perspective needs to be taken. Learning the behavior patterns, i.e., the survey features that influence the willingness to participate into surveys, for respondents and nonrespondents, can help with a more useful assignment of resources to surveys. Striving for an optimal resource allocation can help with reducing the budget overruns.

In the current paper we give a brief overview of the ingredients of adaptive designs such as survey strategies, population characteristics, contact and response probabilities, cost and quality functions and strategy allocation probabilities. Tailoring of design features has been implemented in practice but not necessarily as a result of an optimal balance between quality and costs. Our research aims at optimizing the resource allocation for an adaptive survey design, where the focus is on interview modes, number of time slots and constraints on costs and capacity.

One of the key components of adaptive designs that has not been discussed here is the estimation of input probabilities. A great deal of attention has to be paid to this phase since the optimization part builds upon this input. Issues such as time-dependency, history-dependency, and repeated shifts between interview modes have to be taken into account when estimating the input probabilities.

As mentioned before, the current paper deals with aspects of static adaptive designs. In order to extend to a dynamic setting, additional effort has to be first put into developing techniques to collect, store and utilize paradata such that input parameters and strategy allocation probabilities are updated in real-time. Once these techniques are defined, our method could address a dynamic adaptive design.

Flexibility in addressing various objective functions is another aspect of interest. Recent literature on survey methodology argues that aiming for high response rates influences negatively the bias of the estimators (see, e.g., [7] and [4]). Other quality measures, such as low variation in the group response rates (i.e., representativeness of the respondent sample), have been indicated as more suitable. Currently, the method presented in Section 3 cannot address such a quality function.

Future research aims at tackling these issues in order to develop a model that meets practical needs. Intuitively, taking two survey designs with similar settings, an adaptive design is expected to outperform the classical design since more information becomes available from historical data and the design is tailored such that the group response rate is optimized.

REFERENCES

[1] M.H. Hansen and W.N. Hurwitz. The problem of non-response in sample surveys. *Journal of American Statistical Association*, 41, 1946.

[2] S.G. Heeringa and R.M. Groves. Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(3), 2006.

[3] W.D. Kalsbeek, S.L. Botman, J.T. Massey, and P-W. Liu. Cost-efficiency and the number of allowable call attempts in the National Health Interview Survey. *Journal of Official Statistics*, 10(2):133–152, 1994.

[4] C. Mohl and F. Laflamme. Research and responsive design options for survey data collection at Statistics Canada. In *Proceedings of the American Statistical Association, Survey Research Methods Section*, 2007.

[5] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[6] B. Schouten, M. Calinescu, and A. Luiten. Optimizing quality of response through adaptive survey designs. Working paper.

[7] B. Schouten, F. Cobben, and J. Bethlehem. Indicators for the representativeness of survey response. *Survey Methodology*, 35(1), 2009.

[8] J.R. Wagner. *Adaptive Survey Design to Reduce Nonresponse Bias*. PhD thesis, University of Michigan, 2008.