Using Integer Programming to Determine the Optimal Assignment of Strategies for Mode Switching from Internet to Mail*

John Chesnut
U.S. Census Bureau, Washington, DC 20233

Abstract
With many surveys now embracing adaptive survey design methods, new challenges have arisen in applying optimization methods to meet the requirement to produce automated real-time decisions to adapt survey design strategies across sample cases simultaneously. Chesnut (2013) describes a mode-switching process for switching sample cases from Internet to mail using integer programming with the objective to maximize timeliness while controlling cost, response, and sample representativity. Building on this work, this paper presents the application of an integer programming solution using the OPTMODEL procedure in SAS® to automate the mode-switch decision process. We discuss the use of indicator variables to enable linear representations of our objective function for maximizing timeliness and the constraints for cost, response, and sample representativity in the integer programming environment. While this allows for a complete representation of the optimization problem at hand, the numerous constraints pose a challenge for computing a solution. We discuss alternative linear representations of our constraints using linear approximation methods to enable computed solutions.

Key Words: mode switching, adaptive survey design, optimization, integer programming

1. Introduction

The basic concept of adaptive survey design entails tailoring the survey design to characteristics of the respondent informed by auxiliary frame data or paradata such that trade-offs between survey costs and errors are optimal. Costs could include both tangible and intangible costs such as time. A straightforward approach for determining optimality would be to review historical data and identify time dependent thresholds (the unit of time being days, contacts, etc.) where group-specific tailored survey design strategies would clearly reduce cost and/or error. Using this information, survey managers can establish decision rules prior to data collection that determine the allocation of survey design strategies among sample cases prior to and/or during data collection. As we increase the number of indicators for cost and error, and attempt to address competing priorities such as reliability and nonresponse bias, the optimization problem becomes more complex. In this case, the point of optimality may become less exact or even unknown. The optimization step for the adaptive design process now becomes an execution of "best practices" or a "rule of thumb" approach. As a result, the core component of the adaptive design concept of evaluating cost-error trade-offs is essentially omitted. To remedy this, a mathematical modelling approach may be more effective in accounting for the increasing complexities encountered in an adaptive design problem and provide better solutions or approximations of optimality. Of course, sufficiently representing the

* Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

complexity of the optimization problem may lead to an intractable model also creating a challenge in determining optimality. To illustrate an applied example of using a mathematical modelling approach, Chesnut (2013) developed and simulated an Internet-to-mail mode switch process using data from the 2011 American Community Survey (ACS) Internet tests. Focusing on the optimization component of this framework, we provide further detail on the integer programming model used to solve the problem of maximizing timeliness while constraining for indicators of cost and error.

2. Methodology

2.1 Predicting Internet Response

The larger framework for developing an Internet to mail mode switch process as described in Chesnut (2013) involved using administrative data sources linked to sample cases from the 2011 ACS Internet Tests to predict daily Internet response propensities at the household-level using a discrete time hazards model. Census researchers designed the 2011 Internet tests to test the feasibility of offering an Internet mode of data collection in the ACS (Matthews, et al. 2012 and Tancreto, et al. 2012). Using the daily propensities, we stratified the sample cases into mode switch groups. The eligible mode switch days covered a two-week period beginning with the initial mailout of paper questionnaires for control cases not offered the Internet option and ending with the mailout of a paper questionnaire to nonrespondents.

2.2 Integer Programming

Schouten et al. (2011) construct a mathematical framework to describe adaptive survey designs. They introduce the concept of formulating an optimization problem to enable the decision process for allocating sample cases to survey design strategies in a static manner (prior to data collection) or dynamically (during data collection). The focus of their mathematical model was to create an objective function that maximizes data quality with cost as a constraint or vice versa. In this same spirit, we control for cost and data quality, however we focus on timeliness as our objective.

In our application, we need to determine the optimal day to switch the members of a given mode switch group such that we minimize the average time to nonresponse follow up while meeting constraints on cost and error. Therefore, we will need to model our objective and constraints as a function of the group mode switch days, restricted to the integer values 0,1,…,13. Due to the discrete nature of our decision variables, we will need to model our decision problem using an integer linear programming methodology. Integer programming problems are combinatorial problems and thus are more difficult to solve than traditional linear programming problems (Ignizio and Cavalier, 1994).

Our integer model includes a fixed cost constraint related to mailing a nonresponse followup questionnaire and two error constraints related to bias and level of response. Our cost constraint assumes fixed costs associated with printing, mailing, postage, and data capture. As a control for unit nonresponse bias, we use a proxy measure called the sample representativity indicator (R-Indicator) as an indirect measure of the level of contrast between respondents and nonrespondents (cf. Schouten et al. 2009). Finally, we include constraints that preserve the period of maximum response within each group with the exception of any group where its maximum daily Internet response propensity fails to exceed an assumed threshold. We allowed the solver to switch any such low-Internet-propensity group at the beginning of the two-week period of eligible mode switch days (i.e., on day zero).

2.3 Specifying the Integer Model

Given our available parameters, our goal was to find a solution set of group-based mode switch days that minimize the average group Internet nonresponse follow-up time as defined by

$\bar{y} = (\sum_{g=1}^{G} n_g M_g)/\sum_{g=1}^{G} n_g$ where M_g is the mode switch day and n_g is the sample size for group g, $M_g \in \{0,1,\ldots,13\}$, constrained by

(i) $K\sum_{g=1}^{G}(1-\sum_{t=1}^{M_g} f_g(t))n_g \leq C_o + C$, where K is the cost per Internet nonresponse attributed to mail follow-up, $f_g(t)$ is the predicted response propensity for group g at time t, C_o is the baseline mail follow-up cost if we were to wait until day 13 to follow-up with all Internet nonrespondents, and C is an accepted cost increase

(ii) $M_g \geq t_g^*$ where $f_g(t_g^*)$ is the local maximum for the probability density function for group g at day t_g^*, i.e., $f_g(t_g^*) \geq f_g(t)\ \forall\ t \neq t_g^*,\ t \in \{1,\ldots,13\}$. Note that we relax the constraint, letting $M_g \geq 0$, in the case where the local maximum $f_g(t_g^*)$ fails to exceed a lower bound cutoff f_{LB_cutoff}, i.e., $f_g(t_g^*) \leq f_{LB\_cutoff}$ (e.g, $f_{LB\_cutoff} = 0.01$).

(iii) $\hat{R} \geq \hat{R}_o$ where $\hat{R}_o$ is the level of sample representativity we would achieve if we were to wait until day 13 to follow-up with all Internet nonrespondents.
$\hat{R} = 1 - 2\hat{S}(\hat{\rho}) = 1 - 2\left[\frac{1}{\hat{N}-1}\left(\sum_{g=1}^{G} w_g\left(\hat{\rho}_g - \bar{\hat{\rho}}\right)^2\right)\right]^{\frac{1}{2}}$
where $\hat{\rho}_g = \sum_{t=1}^{M_g} f_g(t)$, $\bar{\hat{\rho}} = \left(\sum_{g=1}^{G} w_g\hat{\rho}_g\right)/\sum_{g=1}^{G} w_g$, $w_g = \sum_{i=1}^{n_g} w_i$,
$\hat{N} = \sum_{g=1}^{G} w_g$, and w_i is the sample design weight.

The mode switch days across groups $\{M_g\}$ are the decision variables in our objective function. Constraint (i) and (iii) include M_g as part of the summation index for calculating the cumulative sum of daily Internet response propensities for a given group. To represent the summation operation as a linear function we define M_g as a summation of dummy indicator variables.

Let $y_{gt} = \begin{cases} 1\ if\ t \leq M_g \\ 0\ otherwise \end{cases}$

Thus, our objective function becomes

$$\bar{y} = (\sum_{g=1}^{G} n_g M_g)/\sum_{g=1}^{G} n_g = \sum_{g=1}^{G}\sum_{t=1}^{13} y_{gt} n_g/\sum_{g=1}^{G} n_g$$

To preserve the properties of the summation index, we need to restrict $(y_{g1}, y_{g2}, \ldots, y_{g13})$ such that $(y_{g1}, y_{g2}, \ldots, y_{g13}) \in \{(0,0,\ldots,0), (1,0,\ldots,0), (1,1,\ldots,0), \ldots, (1,1,\ldots,1)\}$. Therefore, for the g^{th} group we define the following constraints.

$0 \le y_{g1} \le 1$
$0 \le 12y_{g1} - \sum_{t=2}^{13} y_{gt} \le 12$
$0 \le 11y_{g2} - \sum_{t=3}^{13} y_{gt} \le 11$
$\vdots$
$0 \le 2y_{g11} - \sum_{t=12}^{13} y_{gt} \le 2$
$0 \le y_{g12} - y_{g13} \le 1$

In addition, we can write the predicted cost in constraint (i) as follows.

$$K \sum_{g=1}^{G} (1 - \sum_{t=1}^{M_g} f_g(t)) n_g = K \sum_{g=1}^{G} (1 - \sum_{t=1}^{13} y_{gt} f_g(t)) n_g \le C_o + C$$

Furthermore, we can write the requirement to maintain the period of predicted maximum response in constraint (ii) as follows.

$\sum_{t=1}^{13} y_{gt} \ge t_g^*$ where $f_g(t_g^*)$ is the local maximum.

In order to include the nonlinear R-indicator constraint (iii) in our integer programming model, we need to find a way to represent it as a linear constraint. Note that we can re-write (iii) as the following.

$$\hat{R} = 1 - 2\left[\frac{1}{\hat{N}-1} \cdot \left(\sum_{g=1}^{G} w_g (\hat{\rho}_g - \bar{\hat{\rho}})^2\right)\right]^{\frac{1}{2}} \ge \hat{R}_o$$
$$\Leftrightarrow \sum_{g=1}^{G} w_g \hat{\rho}_g^{\,2} - \hat{N}\bar{\hat{\rho}}^2 \le (\hat{N} - 1)(1 - \hat{R}_o)^2/4$$
$$\Leftrightarrow \sum_{g=1}^{G} w_g \hat{\rho}_g^{\,2} - \hat{N}\bar{\hat{\rho}}^2 \le Q^* \quad \text{where } Q^* = (\hat{N} - 1)(1 - \hat{R}_o)^2/4 \tag{1}$$

We have simplified $\hat{R}$ such that we have reduced the nonlinear nature of $\hat{R}$ to two nonlinear components $\hat{\rho}_g^{\,2}$ and $\bar{\hat{\rho}}^2$. We will need to represent these two nonlinear terms as linear functions of our decision variables. Note the following.

$$\hat{N}\bar{\hat{\rho}}^2 = \sum_{g=1}^{G} w_g \left[\left(\sum_{g=1}^{G} w_g \sum_{t=1}^{M_g} f_g(t)\right) / \sum_{g=1}^{G} w_g\right]^2 = \frac{1}{\hat{N}} (w_1\hat{\rho}_1 + \cdots + w_G\hat{\rho}_G)^2$$
$$= \frac{1}{\hat{N}} \big(w_1^{\,2}\hat{\rho}_1^{\,2} + \cdots + w_G^{\,2}\hat{\rho}_G^{\,2} + 2w_1w_2\hat{\rho}_1\hat{\rho}_2 + \cdots + 2w_{G-1}w_G\hat{\rho}_{G-1}\hat{\rho}_G +$$
$$2w_1w_3\hat{\rho}_1\hat{\rho}_3 + \cdots + 2w_1w_G\hat{\rho}_1\hat{\rho}_G\big)$$
$$= \frac{1}{\hat{N}} \left(\sum_{g=1}^{G} w_g^{\,2}\hat{\rho}_g^{\,2} + \sum_{g>h} 2w_g w_h \hat{\rho}_g \hat{\rho}_h\right) \tag{2}$$

From (1) and (2), we observe that the nonlinear nature of the R-indicator can be further isolated to the square terms $\hat{\rho}_g^{\,2}$ and the product terms $\hat{\rho}_g\hat{\rho}_h$, $g > h$.

Note that we can re-write the $\hat{\rho}_g^{\,2}$ terms as

$$\hat{\rho}_g^{\,2} = \left(\sum_{t=1}^{M_g} f_g(t)\right)^2 = \left(\sum_{t=1}^{13} y_{gt} f_g(t)\right)^2$$
$$= \left(y_{g1} f_g(1) + \cdots + y_{gM_g} f_g(M_g) + \cdots + y_{g13} f_g(13)\right)^2$$
$$= \sum_{t=1}^{13} y_{gt}^{\,2} f_g(t)^2 + \sum_{t>u} 2y_{gt} y_{gu} f_g(t) f_g(u)$$

Note that because of the constraints imposed to preserve the properties of the summation index $(y_{g1}, y_{g2}, \dots, y_{g13}) \in \{(0,0,\dots,0), (1,0,\dots,0), (1,1,\dots,0), \dots, (1,1,\dots,1)\}$, we have

$y_{gt}y_{gu} = y_{gt}$ $(t > u)$ and ${y_{gt}}^2 = y_{gt}$. Therefore, we reduce the second order and product terms to provide a linear representation of ${\hat{\rho}_g}^2$.

$$
{\hat{\rho}_g}^2 = \sum_{t=1}^{13} y_{gt} f_g(t)^2 + \sum_{t>u} 2\, y_{gt} f_g(t) f_g(u) \tag{3}
$$

For $g > h$

$$
\begin{aligned}
\hat{\rho}_g \hat{\rho}_h &= \left(\sum_{t=1}^{M_g} f_g(t)\right)\left(\sum_{t=1}^{M_h} f_h(t)\right) = \left(\sum_{t=1}^{13} y_{gt} f_g(t)\right)\left(\sum_{t=1}^{13} y_{ht} f_h(t)\right) \\
&= \left(y_{g1} f_g(1) + \cdots + y_{gM_g} f_g(M_g) + \cdots + y_{g13} f_g(13)\right) \\
&\quad \left(y_{h1} f_h(1) + \cdots + y_{hM_h} f_h(M_h) + \cdots + y_{h13} f_h(13)\right) \\
&= \sum_{t=1}^{13} \sum_{u=1}^{13} y_{gt} y_{hu} f_g(t) f_h(u)
\end{aligned}
$$

Let us introduce the following indicator variables to substitute in for the $y_{gt}y_{hu}$ terms.

$$
z_{gthu} = \begin{cases} 1 \; if \; y_{gt} = 1 \; and \; y_{hu} = 1 \\ 0 \; otherwise \end{cases}
$$

We include additional constraints that force z_{gthu} to take on the value of $y_{gt}y_{hu}$

$$
\begin{aligned}
z_{gthu} &\leq y_{gt} \\
z_{gthu} &\leq y_{hu} \\
z_{gthu} &\geq y_{gt} + y_{hu} - 1
\end{aligned}
$$

Therefore, we have

$$
\hat{\rho}_g \hat{\rho}_h = \sum_{t=1}^{13} \sum_{u=1}^{13} z_{gthu} f_g(t) f_h(u) \tag{4}
$$

Substituting in the linear representations of the nonlinear terms ${\hat{\rho}_g}^2$and $\hat{\rho}_g \hat{\rho}_h$ $(g > h)$ from (3) and (4), we can now re-write (1) as a linear function of our decision variables.

$$
\begin{aligned}
&\sum_{g=1}^{G} w_g {\hat{\rho}_g}^2 - \hat{N}\bar{\hat{\rho}}^2 \\
&= \sum_{g=1}^{G} w_g \left(\sum_{t=1}^{13} y_{gt} f_g(t)^2 + \sum_{t>u} 2\, y_{gt} f_g(t) f_g(u)\right) \\
&\quad - \frac{1}{\hat{N}} \left(\sum_{g=1}^{G} {w_g}^2 \left(\sum_{t=1}^{13} y_{gt} f_g(t) + \sum_{t>u} 2\, y_{gt} f_g(t) f_g(u)\right) + \right. \\
&\qquad \left. \sum_{g>h} 2 w_g w_h \left(\sum_{t=1}^{13} \sum_{u=1}^{13} z_{gthu} f_g(t) f_h(u)\right)\right) \leq Q^*
\end{aligned} \tag{5}
$$

Based on the 667 mode switch groups created from the Internet test sample and the defined constraints our optimization problem has 112,619,617 total constraints. Luenberger (2003) classifies linear programming problems with thousands of variables and constraints as large-scale problems requiring sophisticated code and mainframe computing resources.

To compute a solution for the optimization model we have specified above, we use the OPTMODEL procedure in the SAS Version 9.2 (SAS Institute Inc., 2008). The OPTMODEL procedure uses the OPTMODEL modeling language and serves as a point of access to several mathematical programming solvers in SAS. Given the binary nature of the variables used in our model, the OPTMODEL procedure applies the mixed integer linear programming solver to compute an exact solution for our model. Running our specified model in SAS, we exceed the available memory resources on a Linux server with 23 six-core CPUs and 283 GB of RAM. Given that the computation required to find an exact solution for our optimization model exceeds the capabilities of the available technology, we must settle for an approximate solution.

2.3 Approximating the R-indicator constraint

To find an effective method for computing an approximate solution, we focus on the nonlinear components of our R-indicator constraint – specifically $\bar{\hat{\rho}}^2$. The indicator variables introduced to represent this nonlinear term represent the bulk of the total count of constraints. Pursuing a simple work-around, we explore the use of simple linear approximation techniques to represent the nonlinear components of this constraint as a means of reducing the total number of constraints – specifically eliminating the addition of the indicator variable constraints. Some basic linear approximation methods include linear interpolation, Taylor series approximation, piecewise linear approximation, and ordinary least squares (OLS) regression. Reviewing these different methods, the piecewise linear approximation would provide the best approximations for our nonlinear function, however this is not trivial to implement for multivariate nonseparable functions as in our case. Linear interpolation and first order Taylor series approximations do not provide uniformly good approximations across the range of values for our function, but this is not the case for OLS regression.

Including this additional approximation step essentially converts our optimization problem into a multi-level or hierarchical optimization problem. For example, OLS regression involves minimizing the sum of squares of residuals (the differences between the actual and predicted values).

For the set S of all possible mode switch strategy implementations, we use OLS regression to approximate $\bar{\hat{\rho}}_s^2$ by regressing $\bar{\hat{\rho}}_s^2$ on $\bar{\hat{\rho}}_s$. In other words, we fit the regression line $y_s = \beta_0 + \beta_1 \bar{\hat{\rho}}_s$ such that β_0 and β_1 minimize $\sum\left(\bar{\hat{\rho}}_s^2 - y_s\right)^2$.

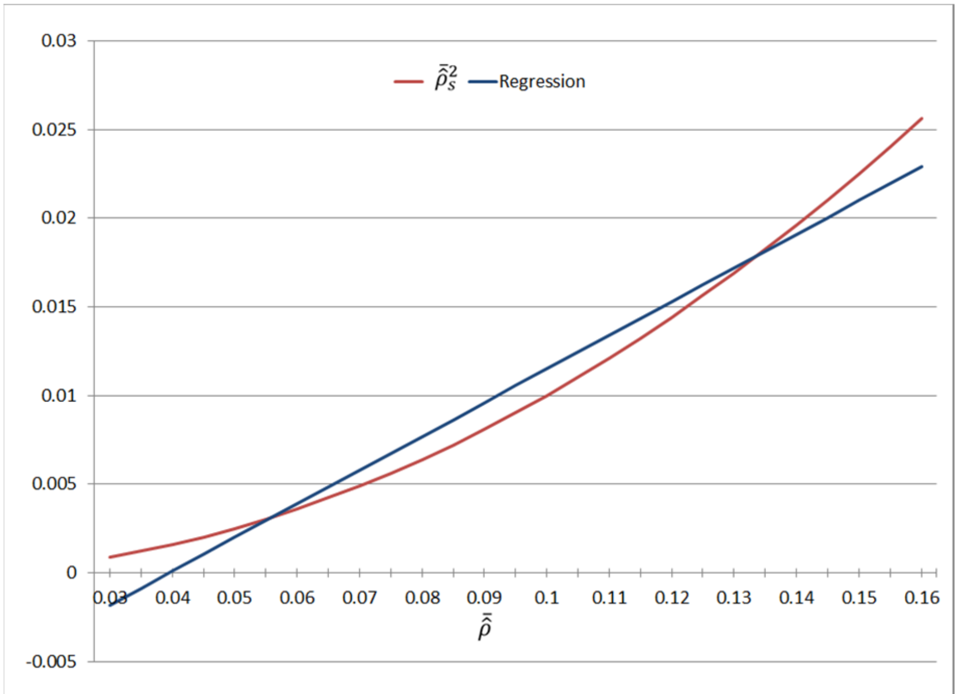
Figure 1 shows the plot of $\bar{\hat{\rho}}_s^2$ and our linear approximation of this function. The curve representing $\bar{\hat{\rho}}_s^2$ appears to be gradual enough that a linear approximation of these values is reasonable. As expected, the regression approach performs well at minimizing the differences between the estimated and actual values of $\bar{\hat{\rho}}_s^2$ (the maximum residual value for the regression approximation was 0.0018 and the average value was 1.8 x 10^{-18}).

Satisfied by these results, we proceed with the regression-based estimates of $\bar{\hat{\rho}}_s^2$ to substitute into the R-indicator constraint. We re-write the R-indicator constraint in (1) as follows.

$$\sum_{g=1}^{G} w_g \hat{\rho}_g{}^2 - \widehat{N}\bar{\hat{\rho}}^2 = \sum_{g=1}^{G} w_g \hat{\rho}_g{}^2 - \widehat{N}\left(\beta_0 + \beta_1 \bar{\hat{\rho}}\right)$$
$$= \sum_{g=1}^{G} w_g \left(\sum_{t=1}^{13} y_{gt} f_g(t)^2 + \sum_{t>u} 2\, y_{gt} y_{gu} f_g(t) f_g(u)\right)$$
$$- \sum_{g=1}^{G} w_g \left(\beta_0 + \beta_1 \frac{\sum_{g=1}^{G}\sum_{t=1}^{13} w_g y_{gt} f_g(t)}{\sum_{g=1}^{G} w_g}\right) \leq Q^*$$

With our linear approximation of constraint (iii), we now have reduced the number of constraints from 112,619,617 to 9,340. While we still consider this a large-scale optimization problem, we now have a more manageable model.

Figure 1. Linear Approximations of $\bar{\bar{\rho}}^2$ Using Ordinary Least Squares Regression

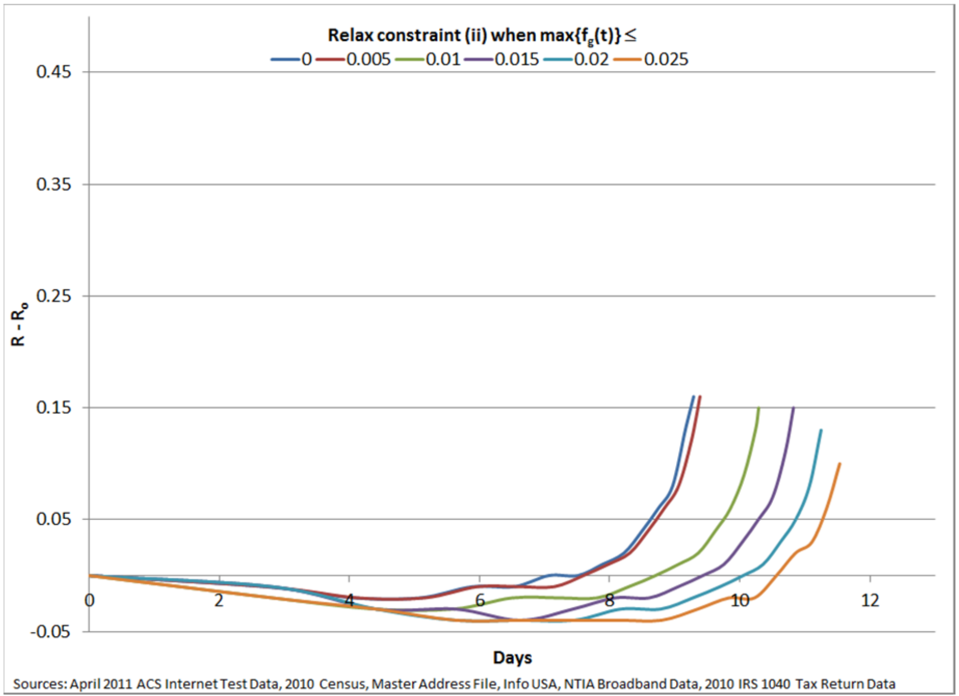


3. Results

3.1 Simulating the Mode Switch Process

Given that we were not able to find an exact solution for optimality that constrains the R-indicator, we want to examine whether the approximation does an adequate job in controlling the sample representativity such that we do not drop substantially below our established baseline level. To serve as a benchmark for comparison, we first simulate our mode switch process without implementing the sample representativity constraint and review the outcomes for the change in R-indicator values relative to our baseline. Figure 2 plots the changes in sample representativity relative to our baseline, $(\hat{R} - \hat{R}_o)$, by the achieved timeliness objective due to mode switching. The minimized average reduced time to nonresponse follow up $(\bar{y})$ is subtracted from the 13 day interval to derive the x-axis values. Each line plot represents a different assumed level of control for relaxing the constraint on preserving a group's period of maximum response. In addition, each data point moving from left to right on a given line plot represent a one percent controlled increase in the cost constraint (i), (1 to 15 percent). We observe that as we progress towards more aggressive mode switching strategies to improve timeliness (i.e., timelier, but more costly), our sample representativity suffers initially. Then depending on our assumptions on when we relax the threshold for preserving the period of maximum response (0.000 to 0.025), the representativity begins to improve (7.5 to 10.5 days for the range of thresholds) to a point where the R-indicator value exceeds that of our baseline value. Clearly, this display of our proxy indicator of the contrast between respondents and nonrespondents at this initial phase of data collection indicates that the contrast is out of control. Note that the multi-mode survey design may compensate for any loss of representativity initially in the later modes of data collection. However, if we could control this outcome earlier in the data collection process we may alleviate some of the burden and expense for the later modes.

Figure 2. The Mode Switching R-indicator Relative to the Baseline by the Average Days Reduced (without the R-indicator Constraint)

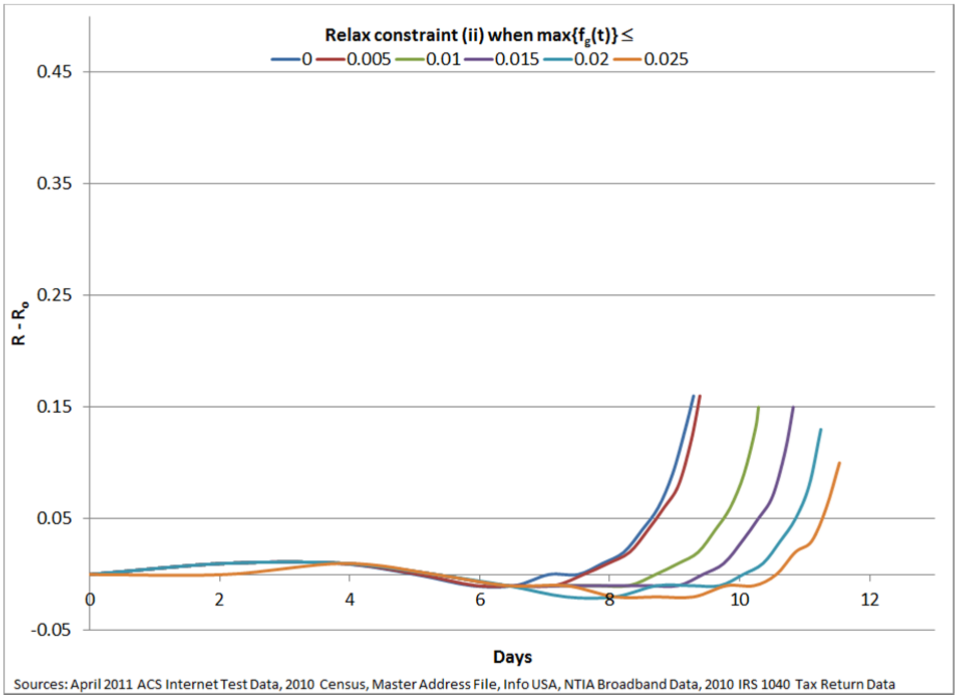


3.2 Implementing the R-indicator Constraint

Next, we simulate our mode-switching process including our approximated R-indicator constraint, and examine the outcome in R-indicator values relative to the baseline. Compared to the line plots in Figure 2, Figure 3 reveals a more controlled level of R-indicator values relative to the baseline. However starting around a 6-day improvement in timeliness, we notice a slight departure from our baseline and then a recovery shortly thereafter. This departure appears to be more sustained for the more liberal thresholds for relaxing the group maximum response constraint (ii). The largest decline relative to the baseline is a value of 0.02 representing only a 2.5 precent decline in sample representativity. Based on these results, we conclude that our solution, while not exact, is well within acceptable tolerances for controlling the level of the R-indicator measure of contrast between respondents and nonrespondents.

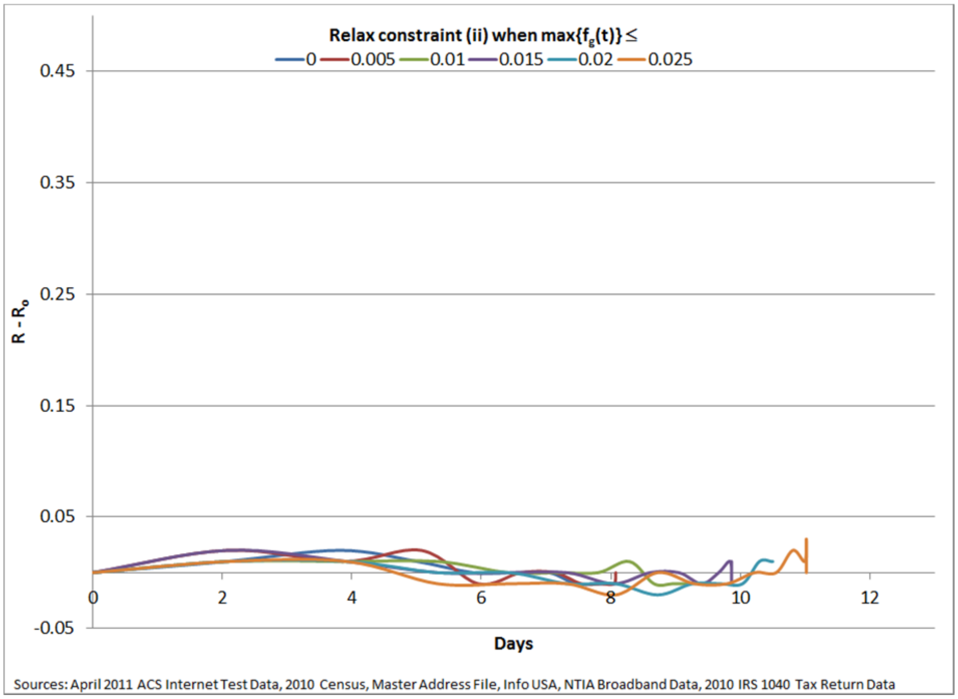
Implementing a lower bound constraint for our R-indicator has allowed the mode switching strategy allocations to produce R-indicator values that approximately follow or exceed the baseline. In Figure 3 we notice that for the more aggressive strategy allocations for improving timeliness the R-indicator dramatically increases relative to the baseline. Given the proxy nature of the R-indicator to measure response bias, we are not entirely confident in using a mode switch allocation that results in extreme swings in the values of this measure. As a result, we implement a more conservative approach, by changing our R-indicator constraint such that we maintain the baseline level across our range of assumption for simulating the mode switch process. Thus, we rewrite constraint (iii) as an equality constraint such that $\hat{R} = \hat{R}_o$.

Figure 3. The Mode Switching R-indicator Relative to the Baseline by the Average Days Reduced (with the Lower Bound R-indicator Constraint)



From Figure 4, we observe the results of our mode switch process with the equality constraint implemented. Clearly, we have a much more controlled outcome in the R-indicator relative to the baseline both bounded from below and above. In other words, we now are approximately maintaining the same level of contrast between respondents and nonrespondent across our range of assumptions. The downside to

Figure 4. The Mode Switching R-indicator Relative to the Baseline by the Average Days Reduced (with the Equality R-indicator Constraint)



implementing this constraint is that we have increased the required processing time by 127 percent as a result of SAS determining a reduced feasible region for our solution set.

4. Conclusion

Through our use of integer programming and the OPTMODEL procedure in the SAS software, we were able to specify a mathematical model for mode switching sample cases from Internet to mail and simulate the mode switch process using the 2011 ACS April Internet Test data. An added complexity to our model included a nonlinear constraint that controlled for the level of sample representativity using the R-indicator measure. This constraint posed a challenge in our attempt to specify this constraint as a linear function of the decision variables. However, by extensive use of indicator variables, we were able to linearize this constraint. While we were able to specify our model to provide an exact solution using this technique, we were not able to compute a solution even with capable computing resources given the intractable number of constraints. As a result, we used a regression-based linear approximation method to approximate the nonlinear constraint to enable a computed solution for meeting the objective of improved timeliness while controlling for cost and error.

Building on the application presented here for mode switching in the self-response modes of data collection, we need to develop a more comprehensive solution and operationalize the concept. For example, we need to assess more accurately and model all costs (tangible and intangible) impacted by mode switching in the self-response modes. In addition, given a fixed period of time for the self-response mode of data collection, we need to also take into account the daily likelihood of mail response for a given household when tailoring the mode switch day such that we are maximizing response in the ‘cheaper modes’. To move beyond the proof of concept phase, we will need to conduct field tests to research how best to operationalize the mode switch process.

By definition, the success of an adaptive survey design process is largely dependent on how effective the decision-making for tailoring survey strategies informed by auxiliary and paradata manages the trade-offs between cost and error. As a result, we view mathematical modelling as the core of the adaptive survey design process. While exact solutions are not always achievable, we demonstrated that approximate solutions can have acceptable results and will have known or approximate optimality rather than relying on methods such as best practices, threshold-based decision rules, or subjective decision-making. In addition, mathematical modelling can account for complex large-scale adaptive survey design problems. Furthermore, in a real-time environment, we can automate the modelling approach as opposed to solutions that introduce lag time by requiring human intervention such as constant monitoring of dashboards.

In general, a more comprehensive adaptive design process than presented here for a multi-mode survey will likely be integer based, nonlinear in nature, and large scale thus posing model tractability issues. To arrive at a comprehensive solution to the adaptive survey design problem, survey methodologists may need to look to the operations research and industrial systems engineering disciplines for assistance.

References

Chesnut, John (2013), "Model-Based Mode of Data Collection Switching from Internet to Mail in the American Community Survey," 2013 JSM Proceedings of the Section on Survey Research Methods, pp 2209-2223.

Ignizio, J. P. and Cavalier, T. M. (1994), *Linear Programming*, Englewood Cliffs, NJ, Prentice Hall.

Luenberger, D. (2003), *Linear and Nonlinear Programming* (2nd ed.), Norwell, MA, Kluwer Academic Publishers.

Matthews, B., Davis, M.C., Tancreto, J., Zelenak, M.F., and Ruiter, M. (2012), "2011 American Community Survey Internet Tests: Results from Second Test in November 2011," 2012 American Community Survey Research and Evaluation Report Memorandum Series #ACS12-RER-21.

SAS Institute Inc. (2008), *SAS/OR 9.2 User's Guide: Mathematical Programming*, Cary, NC: SAS Institute Inc.

Schouten, B., Calinescu, M. and Luiten, M. (2011), "Optimizing Quality of Response through Adaptive Survey Designs," Discussion Paper, Statistics Netherlands.

Schouten, B., Cobben, F. and Bethlehem, J. (2009), "Indicators for the Representativeness of Survey Response," Survey Methodology 35 (1): 101-113.

Tancreto, J., Zelenak, M. F., Davis, M., Ruiter, M., and Matthews, B. (2012), "2011 American Community Survey Internet Tests: Results from First Test in April 2011," 2012 American Community Survey Research and Evaluation Report Memorandum Series #ACS12-RER-13.