MANAGEMENT SCIENCE
Vol. 55, No. 5, May 2009, pp. 727–737
ISSN 0025-1909 | EISSN 1526-5501 | 09 | 5505 | 0727

DOI 10.1287/mnsc.1080.0969

An Optimal Contact Model for Maximizing Online Panel Response Rates

Scott A. Neslin
Tuck School of Business, Dartmouth College, Hanover, New Hampshire 03755,
scott.neslin@dartmouth.edu

Thomas P. Novak
A. Gary Anderson Graduate School of Management, University of California, Riverside,
Riverside, California 92521, tom.novak@ucr.edu

Kenneth R. Baker
Tuck School of Business, Dartmouth College, Hanover, New Hampshire 03755,
kenneth.r.baker@dartmouth.edu

Donna L. Hoffman
A. Gary Anderson Graduate School of Management, University of California, Riverside,
Riverside, California 92521, donna.hoffman@ucr.edu

We develop and test an optimization model for maximizing response rates for online marketing research survey panels. The model consists of (1) a decision tree predictive model that classifies panelists into "states" and forecasts the response rate for panelists in each state and (2) a linear program that specifies how many panelists should be solicited from each state to maximize response rate. The model is forward looking in that it optimizes over a finite horizon during which S studies are to be fielded. It takes into account the desired number of responses for each study, the likely migration pattern of panelists between states as they are invited and respond or do not respond, as well as demographic requirements. The model is implemented using a rolling horizon whereby the optimal solution for S successive studies is derived and implemented for the first study. Then, as results are observed, an optimal solution is derived for the next S studies, and the solution is implemented for the first of these studies, etc. The procedure is field tested and shown to increase response rates significantly compared to the heuristic currently being used by panel management. Further analysis suggests that the improvement was due to the predictive model and that a "greedy algorithm" would have done equally well in the field test. However, further Monte Carlo simulations suggest circumstances under which the model would outperform the greedy algorithm.

Key words: online survey panels; survey response rates; predictive modeling; optimal contact strategy; customer relationship management
History: Received August 29, 2006; accepted November 10, 2008, by Jagmohan S. Raju, marketing. Published online in *Articles in Advance* February 12, 2009.

1. Introduction

Consumer panels have long been a source of respondents for marketing research surveys. Since the explosion of the Internet, marketers have turned to *online* consumer panels (Sudman and Wansink 2002, Carleton 2006, Thornton 2005, Hogg 2003). Online panels promise higher response rates because panelists have precommitted to respond. However, there is evidence that online panels have begun suffering the same response rate problems as traditional methods (Stenbjerre and Laugesen 2005). Davis (2003) reported panel response rates for 2002 as high as 70%, whereas Harris Interactive (2006) more recently reported response rates of 7%–40%. A survey of market researchers suggested that a source of dissatisfaction with online panels is low response rates (Greenfield 2005).

Low response rates raise concern for nonresponse error (Lehmann et al. 1998, pp. 307–309) and imply smaller samples, and therefore larger sampling error. Increasing sample size by inviting more panelists increases incentive costs and risks panelist "burn out." Recruiting more panelists is costly. The key problem is response rates; in fact, legal proceedings scrutinize response rates in determining the validity of survey evidence (Federal Supplement 2002). This suggests a model for maximizing online panel response rates would be desirable. In this paper, we present such a model.

The purpose of this paper is to develop an optimal contact model for maximizing online panel response rates. The model is forward looking in that it takes into account the requirements of future studies in prescribing the sampling plan for the current study. The model also incorporates sample composition requirements, accounts for panel growth, and uses rolling horizon implementation. Overall, the contribution of our paper is as follows:

• We develop and implement a model for increasing online panel response rates.

• To our knowledge, this is the first management science analysis for improving the practice of online panel management.

• We show in a field test that the approach is an improvement over current managerial practice.

• The combination of techniques we use are individually simple but collectively constitute a novel, rigorous, and sophisticated approach to improving response rates.

• We use Monte Carlo simulation to investigate the circumstances under which our approach would improve over a myopic but powerful optimization heuristic, namely, a "greedy" algorithm.

A key aspect of the model is that it accounts for dynamics in two ways. First, average response rates can differ over time, either because individual panelist response rates change or because the process of inviting some panelists and not others sorts panelists into groups, thus identifying high and low responders. Second, the required number of responses can change over time from study to study. To illustrate the importance of this, consider that we have categorized panelists into four groups ("states"): females who responded to their previous invitation, males who responded to their previous invitation, females who did not respond to their previous invitation, and males who did not respond to their previous invitation. The response rates for these groups are 90%, 40%, 35%, and 4%, respectively. The task is to determine which panelists to invite to participate in two studies, the first requiring 100 respondents and the second 400. First, we implement the greedy algorithm, which uses a simple but myopic rule—invite the highest responders available until the requirements are fulfilled. This would yield the results shown in Table 1.

The method invites 2,838 panelists to meet the 500 respondents required, for an overall response rate of 17.6%. What pulls down the response rate is that the method had to invite 1,688 "Male No's" to garner 67 additional respondents to meet the 400-respondent requirement of study 2. A nonmyopic algorithm would take note of the very high response rate among "Female Yes's." None of these females is available for study 1, but by inviting "Female No's" for study 1, we will convert 35% of them to "Female Yes's," and then can exploit this group's 90% response rate. The result is shown in Table 2.

Table 1 Illustration of Greedy Algorithm

		Study 1			Study 2		
State	Response rate	Available	Invite	Expected response	Available	Invite	Expected response
Female Yes	0.90	0	0	0	0	0	0
Male Yes	0.40	500	250	100	350	350	140
Female No	0.35	550	0	0	550	550	193
Male No	0.04	2,000	0	0	2,150	1,688	67
Total		3,050	250	100	3,050	2,588	400

Table 2 Illustration of Nonmyopic Algorithm

		Study 1			Study 2		
State	Response rate	Available	Invite	Expected response	Available	Invite	Expected response
Female Yes	0.90	0	0	0	100	100	90
Male Yes	0.40	500	0	0	500	500	200
Female No	0.35	550	286	100	450	314	110
Male No	0.04	2,000	0	0	2,000	0	0
Total		3,050	286	100	3,050	914	400

The new plan is both efficient and effective: it uses 1,200 invites (286 + 914) to generate the 500 total required respondents, a response rate of 41.7%, much higher than the 17.6% generated by the greedy algorithm. The point is, endemic to planning panelist invitations for online surveys is keeping in mind future as well as current requirements.

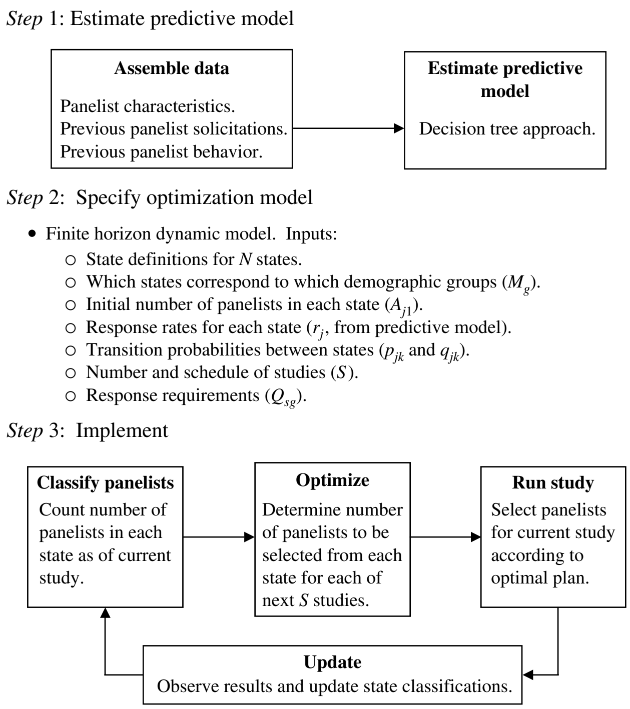
Although the above-listed contributions are important, it is appropriate to acknowledge several limitations up front to illustrate the complexity of issues. First, the superior results of our approach were potentially due both to the predictive model that underlies the optimization, as well as to the dynamics of the optimization. In our particular case, the superior results of our field test were found to be due to the predictive model, rather than the dynamics of the optimization. We are able to examine this issue through Monte Carlo simulation. Second, the model maximizes *expected* response rates and does not explicitly account for uncertainty. This could be addressed by using the Monte Carlo simulation to construct confidence intervals for the response rate. Third, the objective to maximize response rates is subject to sample size goals or "requirements." Due to uncertainty, we may exceed the requirements or not meet them. A Monte Carlo simulation again could be used a priori to quantify the chances of exceeding or not meeting requirements and if necessary fine tune the number of panelists contacted. Finally, in our model, we optimize overall response rates over several studies and do not try to equalize response rates across studies. This could be accomplished most practically by weighting the objective function by study. We encourage further field work to address these important issues.

2. The Model

2.1. Overview

The model's foundation is similar to catalog planning models developed by Bitran and Mondschein (1996),

Figure 1 Overview of Optimal Contact Model



Elsner et al. (2003), and Simester et al. (2006). It relies on a predictive model and an optimization model. The predictive model defines "states" for classifying panelists and forecasts the response rate if panelists in a particular state are invited to participate in a study. The optimization is a linear program that yields a plan of panelist invitations over a series of S studies that maximizes the overall response rate. An overview of the model is in Figure 1.

First, to estimate the predictive model, we assemble data depicting the history of panelist invitations and responses. The data include the "dependent variable," showing whether the panelist responded to an invitation, and a series of potential predictor variables. We utilize a decision tree predictive model because this fits the strategy of classifying panelists into states (see Simester et al. 2006, Blattberg et al. 2008). The "end nodes" of the tree define the states and expected response probability of a respondent in that state. One end node might be "panelists who are female and were never invited to participate in a study." Panelists in this state might have a 20% chance of responding to a solicitation.

Second, we specify the optimization model. The optimization assumes that the schedule for the next S studies is known, and the panel manager wishes to achieve a target number of responses for each study, possibly with representation requirements by demographic group. The optimization is a linear program that minimizes the number of solicitations to achieve the participation levels required for each of the studies in the upcoming S-study horizon.

Third, we use a "rolling horizon" approach to implement the model. Assume the horizon is four studies. We start with the initial numbers of panelists in each state. We optimize over studies 1–4, and implement the plan for study 1. We then observe the results, update our state classifications, and rerun the model for studies 2–5, implementing what that optimization tells us to do for study 2. We then observe the results for study 2, update our state classifications, and rerun the model for studies 3–6, etc. To our knowledge, rolling horizons have not been applied in marketing.[1] The concept is well known in the operations management literature (Baker 1977, McClain and Thomas 1977, Chand et al. 2002). The advantage of rolling horizons is ease of implementation.

2.2. Detail of the Optimization Model

The input parameters for the model are shown in Table 3. The transition probabilities—p_{jk} and q_{jk}—are derived from the definition of states (D_j) and the predicted response rates (r_j). For example, state 1 might be male panelists who were never invited to a study, state 2 might be male panelists who were invited to a study but did not respond, and state 3 might be defined as male panelists who were invited to a study and did respond. Assume that $r_1 = 0.3$, $r_2 = 0.1$, and $r_3 = 0.8$. Then, we would have

$$\begin{array}{llllll} p_{11}=0 & p_{12}=0.7 & p_{13}=0.3 & q_{11}=1 & q_{12}=0 & q_{13}=0 \\ p_{21}=0 & p_{22}=0.9 & p_{23}=0.1 & q_{21}=0 & q_{22}=1 & q_{23}=0 \\ p_{31}=0 & p_{32}=0.2 & p_{33}=0.8 & q_{31}=0 & q_{32}=0 & q_{33}=1. \end{array}$$

The decision variables for the linear program are as follows:

X_{js} = Number of panelists from state j invited to participate in study s.

A_{js} = Number of panelists in state j available for participation in study s.

Given these decisions, the ps and qs determine the number of panelists who will be available in various states to participate in future studies. The optimization is then

$$\text{Minimize} \quad \sum_{s=1}^{S}\sum_{j=1}^{N} X_{js}$$

$$\text{subject to} \quad X_{js} \le A_{js} \quad j=1,\ldots,N;\ s=1,\ldots,S; \tag{1}$$

$$\sum_{j\in M_g} r_j X_{jt} \ge Q_{sg} \qquad g=1,\ldots,G;\ s=1,\ldots,S; \tag{2}$$

$$A_{ks}=\sum_{j=1}^{N} p_{jk}X_{j,s-1}+\sum_{j=1}^{N} q_{jk}(A_{j,s-1}-X_{j,s-1})+R_{ks} \qquad k=1,\ldots,N;\ s=2,\ldots,S. \tag{3}$$

[1] Note that Elsner et al. (2003) refer to Elsner (2002) as proposing rolling horizons for catalogs.

Table 3 Definitions of Terms

Input parameters

N = Number of states in which panelists can be classified.
S = Number of studies in the time horizon.
G = Number of demographic groups ($G \leq N$).
$j, k = 1, \ldots, N$ index states.
$s = 1, \ldots, S$ indexes studies.
$g = 1, \ldots, G$ indexes demographic groups.
D_j = Definition of state j, i.e., the values of the predictor variables that define a particular state.
M_g = Set of states whose panelists are in demographic group g.
Q_{sg} = Target number of responses desired from demographic group g for study s.
r_j = Response rate for panelists in state j.
p_{jk} = Proportion of panelists in state j who migrate to state k if they are invited to participate in a study.
q_{jk} = Proportion of panelists in state j who migrate to state k if they are not invited to participate in a study.
A_{j1} = Number of panelists in state j available for the first study.
R_{js} = Number of new panelists who exogenously enter state j in time to participate in study s.

Decision variables

X_{js} = Number of invitations from state j to participate in study s.
A_{js} = Number of panelists in state j available for study s ($s = 2, \ldots, S$).

The objective is to minimize the number of invitations over the horizon. This is equivalent to maximizing the overall response rate across studies. Constraint (1) ensures we cannot invite more panelists from a given state, for a given study, than are available. Constraint (2) imposes the target number of responses for demographic groups and studies. Constraint (3) calculates the number of panelists available for each study from each state given the invitation plan (X), the transition probabilities (p and q), and the arrival of new panelists (R). Availabilities (A_{js}) are provided as input for $s = 1$.

3. Application

3.1. Setting and Data

The application is for an online panel maintained at a university research center that provides respondents for academic research studies. Response rates had taken a precipitous plunge (Figure 2), and management wished to increase response rates. Before study 38, management sent out a "request for confirmation" e-mail to verify the e-mail addresses of all panel members. A positive response should be a good predictor of whether the panelist would respond to invitations to participate in future studies.

Figure 2 Online Panel Response Rates for Recent Studies

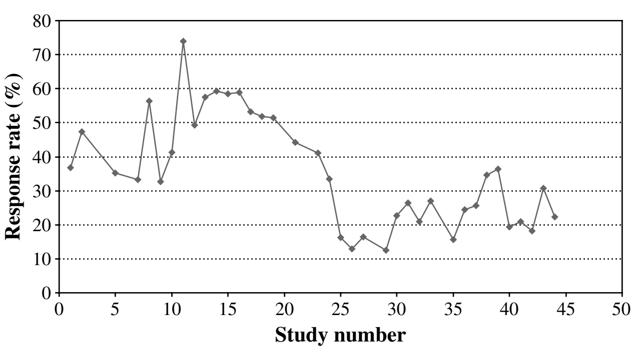


Table 4 Variables Available for Predictive Model

Panelist characteristics

1. Confirmation e-mail status (OK, Not OK, Not asked)
2. Year joined panel
3. Agreed to opt-in e-mail for other solicitations
4. Gender
5. Education
6. Country of origin
7. Acquisition source
8. Age

Previous invitations and behavior

9. Days between joining panel and first invitation
10. Days between 1st and 2nd invitation
11. Days between 2nd and 3rd invitation
12. Days between 3rd and 4th invitation
13. Response to previous invitation (Yes, no, or never been invited)

Invitation-specific

14. Days between previous invitation or joining panel and this invitation
15. Study number
16. Date invitation was mailed
17. Days between previous invitation and this invitation

Notes. Each row of data corresponds to a given panelist's invitation to participate in a particular study. If a panelist was invited to more than one study, there would be a row for each invitation.

Management assembled a database of 38,212 observations, with an observation corresponding to a given panel member's invitation to participate in a particular study. Table 4 shows the variables. Beginning with study 38, panel management consistently used a simple heuristic: select (1) those who responded "yes" to the confirmation e-mail and (2) those who joined the panel in the last three months. Thus, we decided to start the database at study 38. This left 12,129 observations. Before estimating the model, panel management stipulated that they wished to specify response quotas at the gender level. This meant that $G = 2$, with $g = 1$ signifying male and $g = 2$ signifying female.

When estimating the decision tree, a variable included on one branch needed to be included on the other branches as well. Consider the case that the decision tree split first on time since becoming a member of the panel, with the dividing point "the last 50 days" versus "more than 50 days ago." Assume the last-50-days branch then split on gender. We then also have to include gender as a variable on the more-than-50-days-ago branch of the tree. The reason is that, over time, a male panelist who joined during the last 50 days would eventually transition to the more-than-50-days-ago branch of the tree. To know how this would affect response rate, we would need the gender variable on that side of the tree as well. To test

Table 5 Predictive Model Results and States

State	Days between last invite or join and current study	Did panelist complete previous invite?	Panelist response to request to confirm status	Gender	Sample size	Raw response rate	Pooled response rate	Next state if invited and respond[a]	Next state if invited and do not respond[a]	Next state if not invited[a]
1	0–61	Yes	OK	M	41	0.5122	0.6328	1	7	19
2	0–61	Yes	OK	F	22	0.7273	0.6328	2	8	20
3	0–61	Yes	Not OK	M				3	9	21
4	0–61	Yes	Not OK	F				4	10	22
5	0–61	Yes	Not asked	M	38	0.6579	0.6328	5	11	23
6	0–61	Yes	Not asked	F	27	0.7037	0.6328	6	12	24
7	0–61	No	OK	M	151	0.1523	0.1731	1	7	25
8	0–61	No	OK	F	132	0.1515	0.1731	2	8	26
9	0–61	No	Not OK	M				3	9	27
10	0–61	No	Not OK	F				4	10	28
11	0–61	No	Not asked	M	83	0.1807	0.1731	5	11	29
12	0–61	No	Not asked	F	50	0.2800	0.1731	6	12	30
13	0–61	No invite	OK	M				1	7	31
14	0–61	No invite	OK	F				2	8	32
15	0–61	No invite	Not OK	M				3	9	33
16	0–61	No invite	Not OK	F				4	10	34
17	0–61	No invite	Not asked	M	1,519	0.3502	0.3663	5	11	35
18	0–61	No invite	Not asked	F	1,937	0.3789	0.3663	6	12	36
19	62+	Yes	OK	M	392	0.4133	0.4133	1	7	19
20	62+	Yes	OK	F	496	0.4698	0.4698	2	8	20
21	62+	Yes	Not OK	M	35	0.0286	0.0286	3	9	21
22	62+	Yes	Not OK	F	51	0.1569	0.1569	4	10	22
23	62+	Yes	Not asked	M	3	0.0000	0.0000	5	11	23
24	62+	Yes	Not asked	F	1	0.0000	0.0000	6	12	24
25	62+	No	OK	M	400	0.1825	0.1825	1	7	25
26	62+	No	OK	F	420	0.2476	0.2476	2	8	26
27	62+	No	Not OK	M	101	0.0198	0.0198	3	9	27
28	62+	No	Not OK	F	109	0.0825	0.0826	4	10	28
29	62+	No	Not asked	M	2	0.0000	0.0000	5	11	29
30	62+	No	Not asked	F	2	0.0000	0.0000	6	12	30
31	62+	No invite	OK	M	1,758	0.2378	0.2356	1	7	31
32	62+	No invite	OK	F	2,868	0.2313	0.2356	2	8	32
33	62+	No invite	Not OK	M	323	0.0310	0.0310	3	9	33
34	62+	No invite	Not OK	F	335	0.0627	0.0627	4	10	34
35	62+	No invite	Not asked	M	362	0.1381	0.1381	5	11	35
36	62+	No invite	Not asked	F	471	0.2059	0.2059	6	12	36

[a]These are the transitions used for the optimization. They are based on the assumption that (1) the time between any two studies is less than 61 days (so that any invited panelist moves to states 1–18 for the current study) and (2) the time between any first and third study is more than 61 days (so that any noninvited panelists move to states 19–36). In the actual field test, the exact timing was 29 days between studies 1 and 2, 28 days between studies 2 and 3, and 29 days between studies 3 and 4. So there was a slight discrepancy between how the field test was actually run and the second assumption made by the optimization. As a result, however, noninvited panelists often stayed in states 1–18, contrary to what was being assumed by the optimization. See Table A.1 in the online appendix.

the impact of this "symmetry" requirement, we investigated three models:[2]

• Model 1 allowed all potential variables in Table 4 to be included in the tree, and allowed the tree to be as deep as the program cutoffs deemed statistically significant. Symmetry was not imposed.

• Model 2 allowed all potential variables as in Model 1, but limited the tree to four levels. Symmetry was not imposed.

• Model 3 used only variables that appeared in the first two levels of the tree, limited the tree to four levels, and imposed the symmetry constraint.

From the variables entering the tree at the first two levels, it was clear that the most important variables were *confirmation status, response to previous invitation, days between previous invitation or joining the panel and the current invitation*, and *gender*. This was fortuitous because we wanted *gender* to be in the model anyway. To determine how much predictive ability we lost by constraining the model to four levels and symmetry, we compared the top decile "lift" for the three models (e.g., see Neslin et al. 2006). The lifts were 1.796, 1.792, and 1.679 for the three models, respectively. This shows that we lost a little bit of predictive power, but not much, in constraining the predictive model as we did.

Table 5 shows Model 3 in tabular form. "*Days between last invite or join and current invitation*" was split at 61 days. A panelist would be placed in the ≤61 days classification if he or she either joined the panel within the last 61 days, or was last solicited

[2] Decision trees were estimated using the chi-square automatic interaction detection (CHAID) procedure in AnswerTree®, distributed by SPSS.

for a study within the last 61 days. Generally, the $\leq$61 days group has a higher response rate. The "*did panelist complete previous invited*?" variable was split into "yes," "no," or "not invited." Generally, "yes's" have higher response rates. For the *confirmation status* variable, "OK's" clearly have higher response rates. Finally, women tend to have higher response rates than men. The 36 end nodes mean 36 possible states for each panelist ($N = 36$).

Certain states have no sample or a very small sample because certain combinations never or rarely occurred in the data. We therefore pooled certain states. For example, we pooled states 1, 2, 5, and 6 so that the predicted response rate was 63.28% for any panelist in one of those states. Pooling was based on statistical tests of whether the raw response rates for the potentially pooled states were equal. However, there were still some states for which we could not make a prediction.

3.2. The Optimization Model and Its Recommended Plan

Table 5 allowed us to calculate transition probabilities for the optimization model. For example, if a panelist in state 1 is invited but does not respond, he migrates to state 7. The probability this occurs is $q_{1,7} = 1 - 0.6328 = 0.3672$. The target number of responses for each study (Q_{sg}) was 200 males and 200 females, for a total of 400 desired responses per study. The growth rate was assumed to be 100 males and 200 females per month.[3] This defined R_{js} in Equation (3). Panel management allocated 33,000 of its 68,329 panelists to the field test, divided into three equal groups according to whether they were to be managed under the optimization model, management's current invitation heuristic, or random selection.

Table 6 shows the optimization plan derived using the above information. We start with 11,000 panelists, most of whom were "old" in that they either had joined or last been solicited >61 days ago. The strategy of the optimal solution is intuitive. It (1) develops a pool of recently solicited panelists who responded and whose confirmation status is either "OK" or "not asked" (states 1, 2, 5, or 6), and (2) because their predicted response rate is 63.28%, uses them to increase response in future studies.

The result was that the optimal plan predicted response rates of 32.3%, 53.3%, 53.7%, and 47.8% across the four studies.[4] The plan in Table 6 was used to solicit for the first study. Our *expectation* was that our actions for study 1 would create a sizable pool of state 1 and 2 panelists, and combined with the growth in the panel (this produced more panelists in high-response states 17 and 18), we would achieve higher response rates for study 2. This rosy picture needed, of course, to be confirmed by the field test.

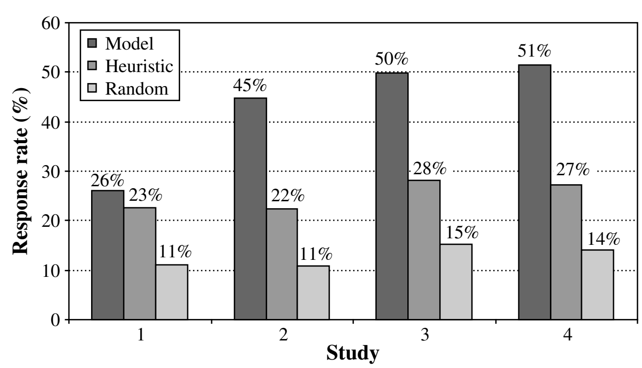
4. The Field Test

The field test compared the optimization model plan to two benchmarks: random selection from the pool of all available panel members and the management heuristic. The heuristic created two groups: (1) those who responded "yes" to the confirmation e-mail and (2) those who joined the panel in the last three months. The groups were merged, and then panelists were randomly selected with the restriction of obtaining 50% women and 50% men (corresponding to the requirement of equal numbers of male and female respondents). From the predictive model, we knew that in fact e-mail confirmation and joining the panel in the last few months were strong positive indicators of response.

To facilitate comparisons, we required each method to invite the same number of panelists to each study. That number was determined by the optimal number prescribed by the model. This did not hamper the heuristic and random approaches in any way because they had a large enough pool.

Figure 3 shows the response rates for the field test. Things went much according to plan. The model performed best for study 1, but not too much better than the heuristic. This is because the model was developing its pool of high-response panelists (states 1 and 2). The payoff came in studies 2–4, where the optimization plans achieved roughly double the response rate of the heuristic. The overall response rate for the model was 41.2%. The detailed results by state and study are available in Tables A.1–A.3 in the online

Figure 3 **Field Test Response Rate Results: Model, Heuristic, and Random Methods**



[3] Management was certain they could add 300 new panelists per month for the field test.

[4] Note that in our specific application, the optimal solution is not unique. When faced with multiple optima, we can generate a few equivalent plans and select the one preferred for other managerial considerations.

Table 6 Initial Optimization Plan

States					Predicted response	Study 1			Study 2			Study 3			Study 4		
State	Days	Prev.	Conf.	Gender	Rate	Avail.	Inv.	Exp. resp.	Avail.	Inv.	Exp. resp.	Avail.	Inv.	Exp. resp.	Avail.	Inv.	Exp. resp.
1	0–61	Yes	OK	M	0.6328	0			187	187	118	155	155	98	135	135	85
2	0–61	Yes	OK	F	0.6328	0			200	200	127	191	191	121	194	194	123
3	0–61	Yes	Not OK	M		0			0			0			0		
4	0–61	Yes	Not OK	F		0			0			0			0		
5	0–61	Yes	No ask	M	0.6328	0			13	13	8	45	45	28	65	65	41
6	0–61	Yes	No ask	F	0.6328	0			0			9	9	6	6	6	4
7	0–61	No	OK	M	0.1731	0			589			121			109		
8	0–61	No	OK	F	0.1731	0			226			146			153		
9	0–61	No	Not OK	M		0			0			0			0		
10	0–61	No	Not OK	F		0			0			0			0		
11	0–61	No	No ask	M	0.1731	0			23			68			80		
12	0–61	No	No ask	F	0.1731	0			0			16			3		
13	0–61	Noinv.	OK	M		0			0			0			0		
14	0–61	Noinv.	OK	F		0			0			0			0		
15	0–61	Noinv.	Not OK	M		0			0			0			0		
16	0–61	Noinv.	Not OK	F		0			0			0			0		
17	0–61	Noinv.	No ask	M	0.3663	36	36	13	100	100	37	100	100	37	100	100	37
18	0–61	Noinv.	No ask	F	0.3663	63			200	25	9	200			200		
19	62+	Yes	OK	M	0.4133	374	196	81	178	89	37	89	89	37	0		
20	62+	Yes	OK	F	0.4698	875	426	200	449	137	64	313	156	73	156	156	73
21	62+	Yes	Not OK	M	0.0286	306			306			306			306		
22	62+	Yes	Not OK	F	0.1569	741			741			741			741		
23	62+	Yes	No ask	M	0.0000	158			158			158			158		
24	62+	Yes	No ask	F	0.0000	242			242			242			242		
25	62+	No	OK	M	0.1825	653	580	106	73			662			783	106	19
26	62+	No	OK	F	0.2476	1,203			1,203			1,429			1,575		
27	62+	No	Not OK	M	0.0198	1,034			1,034			1,034			1,034		
28	62+	No	Not OK	F	0.0826	1,806			1,806			1,806			1,806		
29	62+	No	No ask	M	0.0000	502			502			525			593		
30	62+	No	No ask	F	0.0000	774			774			774			790		
31	62+	Noinv.	OK	M	0.2356	74			74			74			74	74	17
32	62+	Noinv.	OK	F	0.2356	181			181			181			181		
33	62+	Noinv.	Not OK	M	0.0310	463			463			463			463		
34	62+	Noinv.	Not OK	F	0.0627	744			744			744			744		
35	62+	Noinv.	No ask	M	0.1381	298			298			298			298		
36	62+	Noinv.	No ask	F	0.2059	473			536			711			911		
Total						11,000	1,238	400	11,300	751	400	11,600	745	400	11,900	836	400
Total men							812	200		389	200		389	200		480	200
Total women							426	200		362	200		356	200		356	200
Expected response rate (%)								32.3			53.3			53.7			47.8

Note. Prev., previous response; noinv., never previously invited; conf., response to confirmation e-mail; no ask, was not sent confirmation e-mail; avail., number available in each state; inv., number invited from each state; exp. resp., number of expected responses from each state.

appendix (provided in the e-companion).[5] For each study, we also compared (1) the characteristics of the respondents and (2) the mean values of their responses. These numbers do not differ significantly across approaches. Although the differences were not significant, we did note that the model achieved a sample closer to the desired 50–50 male–female split (online appendix, Tables B and C). These findings do not eliminate the possibility of frame or nonresponse bias. Management would need to compare the demographics of panel members, panel responders, and the general public to monitor this issue.

[5] An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Overall, the model yielded a higher response rate and the substantive results of each study did not differ. The benefit is that the sample size was appreciably higher because the response rate was higher. This would yield smaller standard errors and more powerful statistical tests (online appendix, Table E). Alternatively, we could have used the optimal contact model to invite fewer panelists and achieve roughly the same sample size as management's heuristic.

5. Analysis of Panelist Response Rates

As noted in the Introduction, an important dynamic presumed in our analysis is that average panelist

Table 7 Analysis of Panelist Response Rates

Panel A

State	Predicted response rate for each state (Table 5) (%)	Actual response for each state: Response rate (%)	Actual response for each state: Sample size	Predicted response rate next study | respond and invited (Table 5) (%)	Actual response rate next study | respond and invited: Response rate (%)	Actual response rate next study | respond and invited: Sample size	Predicted response rate next study | did not respond and invited (Table 5) (%)	Actual response rate next study | did not respond and invited: Response rate (%)	Actual response rate next study | did not respond and invited: Sample size
7	17.3	(d)	(d)	63.3	50	2	17.3	15.4	26
17	36.6	40.0	25	63.3	(c)	0	17.3	0	1
18	36.6	40.5	37	63.3	(c)	0	17.3	0	1
19	41.3	35.6	135	63.3	100.0	8	17.3	20.0	10
20	47.0	33.1	175	63.3	100.0	5	17.3	50.0	6
21	2.9	8.7	92	(a)	(c)	0	(a)	0	11
22	15.7	11.0	146	(a)	(c)	0	(a)	0	7
23	(b)	27.5	51	63.3	100.0	1	17.3	0	5
24	(b)	37.3	51	63.3	100.0	1	17.3	0	2
25	18.3	10.6	218	63.3	100.0	3	17.3	3.6	28
26	24.8	15.1	199	63.3	50.0	2	17.3	0	7
27	2.0	2.1	330	(a)	100.0	0	(a)	0	28
28	8.3	1.8	342	(a)	(c)	0	(a)	0	20
29	(b)	9.0	178	63.3	66.7	3	17.3	4.8	21
30	(b)	6.4	140	63.3	(c)	0	17.3	0	11
31	23.6	20.0	20	63.3	100.0	1	17.3	10.0	10
32	23.6	12.5	32	63.3	(c)	0	17.3	0	3
33	3.1	2.1	146	(a)	(c)	0	(a)	0	13
34	6.3	2.3	133	(a)	(c)	0	(a)	0	8
35	13.8	9.8	123	63.3	100.0	1	17.3	0	11
36	20.6	13.6	118	63.3	0.0	1	17.3	0	5

Panel B

State	Predicted response rate for each state (%)	Actual response rate for each state (%)	Predicted direction of change | responded and invited to next study	Actual increase or decrease	Predicted direction of change | did not respond and invited to next study	Actual increase or decrease
7	17.3	(d)	Increase	Increase	No change	(e)
17	36.6	40.0	Increase	(c)	Decrease	Decrease
18	36.6	40.5	Increase	(c)	Decrease	Decrease
19	41.3	35.6	Increase	Increase	Decrease	Decrease
20	47.0	33.1	Increase	Increase	Decrease	Decrease
25	18.3	10.6	Increase	Increase	Decrease	Decrease
26	24.8	15.1	Increase	Increase	Decrease	Decrease
31	23.6	20.0	Increase	Increase	Decrease	Decrease
35	13.8	9.8	Increase	Increase	Increase	Decrease
36	20.6	13.6	Increase	Decrease	Decrease	(f)

(a) Cannot predict response rate because panelist would move to a cell for which we had zero observations in the calibration data.
(b) Cannot predict response rate due to insufficient sample size ($n = 3$ for state 23; 1 for state 24; 2 for state 29; 2 for state 30).
(c) Cannot calculate response rate because sample size = 0.
(d) Did not calculate initial response rate because panelist could be in that state either because they were invited or had been in that state for a previous study.
(e) Predict no change, so not relevant to assess increase or decrease.
(f) No definitive answer—depends on whether predicted or actual response rate is used as base.

response rates can change over time. We now investigate whether in fact (1) average panelist response rates differ as they move from state to state over time and (2) these rates differ in the manner we predict. We use the random sample condition from the field test for this analysis. The sample sizes become small because less than a tenth of panelists are invited for any single study (Table A.2, online appendix). However, our strategy is to find the states for which we can predict and then observe the direction in which these response rates should change. Though the sample sizes are often small, we examine whether the *pattern* of increases in decreases in response rates is in accord with our predictions.

We first had to eliminate states 1–6 and 8–16 from this analysis because either we did not have any observations on these states when we estimated the predictive model (states 3, 4, 9, 10, and 13–16), or we had no observations for the state transitions where we would predict a change in response rate (states 1, 2, 5, 6, 8, 11, and 12). We were however able to assemble Table 7.

Consider state 19 in Table 7 (panel A). The predictive model predicts a 41.3% response rate for panelists in this state. For panelists who respond, the model predicts an *increased* response rate, 63.3%, if they are also invited to the next study. For panelists who do not respond, the model predicts a *lower* response rate, 17.3%, if they are also invited to the next study. Table 7 (panel A) shows that the actual response rate for panelists in state 19 was 35.6%, the response rate for panelists in state 19 who responded and were invited to the next study was 100% (8 out of 8), and for those who did not respond yet were invited to the next study it was 20% (2 out of 10). So the response rates differed longitudinally, and in the predicted direction.

Our goal was to do a similar analysis for the other states. However, this was not possible for all the states in Table 7 (panel A). We eliminated from further analysis states 23, 24, 29, and 30 because there were only one to three observations in the predictive model for those states; we eliminated states 21, 22, 29, 30, 33, and 34 because panelists invited from these states moved to states for which the sample size in our predictive model was too small to forecast response rates; we eliminated state 32 because either panelists invited from this state moved to a state for which the sample size in our predictive model was too small, or because they would move to a state for which there is no definitive prediction of increase or decrease (the predictive model said decrease, whereas the actual observed response rate said increase).

Fortunately, we are able to observe "clean" results for the 10 states shown in Table 7 (panel B). In total, we make 16 predictions regarding the direction of change in response rates where we can observe what actually happened. These predictions are correct 14 times, and incorrect two times. A two-sided sign test indicates that the probability of observing a result this extreme, if we were only guessing (50% success rate), is 0.0042 (http://graphpad.com/quickcalcs/binomial1.cfm).

Overall, there is a clear pattern. We were able to predict correctly 14 out of 16 times the direction in which average response rates change as customers are invited to a study and progress to another state because of that invitation. These results suggest that (1) average response rates differ over time as customers move from state to state, and (2) we predict the direction of these changes with consistent accuracy. This provides support that the predictive model captures the dynamics of how response rates change.

6. Monte Carlo Simulation

As mentioned at the beginning of the paper, our model optimizes expected response rates and does not take into account uncertainty. We conduct a Monte Carlo simulation to generate the distribution of potential outcomes for a given application of the model. The simulation takes into account two sources of uncertainty: (1) the predicted response rates (Table 5) are based on a finite sample size and hence subject to sampling error; (2) we invite a finite number from each state, and hence the actual number of respondents follows a binomial distribution. The simulation runs as follows: First, we address the first source of uncertainty by drawing an actual response rate for each state i from a normal distribution with mean π_i and variance $\pi_i(1-\pi_i)/n_i$, where n_i is the pooled sample size used in Table 5.[6] Define p_i as the randomly generated response rate for state i. Second, we use the p_is to run the optimization, and assume we invite x_i panelists from state i for the first study. Third, we account for the second source of uncertainty by drawing the actual number of responses from a binomial distribution with parameters x_i and p_i. This yields the actual number of panelists in each state as we plan for the second study. Fourth, we use the p_is to run the optimization for the second study, generating the x_is for that study. We then repeat the process for the third and fourth studies, using the rolling horizon method. Fifth, we use 1,000 replications of steps 1–4 to derive the distribution of response rates.

The first important result is that, using the field test demands of 400 for each study (400-400-400-400), the mean response rate across the four studies is 43.8% with a standard deviation of 2.9%. Because the distribution across replications was approximately normal, this means a prediction for the overall response rate would be 43.8% ± 5.8%. The actual overall response rate, as mentioned earlier, was 41.2%. This means the Monte Carlo simulation accurately predicts what actually occurred in the field test.

We conducted simulations to compare our model to the greedy algorithm, which is a formidable competitor to systematic optimizations (see Ansari and Mela 2003). The greedy algorithm selects panelists from states that have the highest response rates, until it predicts it has selected enough panelists to satisfy the sample requirements (Table D in the online appendix). It does not take into account future demands. The field test, with the same sample requirements for each study, did not impose a challenge on the dynamic aspect of our model. Hence, it was not surprising that the Monte Carlo simulation found that the greedy algorithm would have outperformed the model (44.9%

[6] We assume the pooled response rates, which are the response rates we used for the optimization in the field test, are unbiased and hence use them for the true response rate π_i. The important point is that we incorporate sampling error in estimating π_i by using its sampling distribution.

response rate versus 43.8% for our model).[7] This finding shows that the gain over management practice in the field test was due to the power of the predictive model, not to dynamics in the optimization. However, we simulated a second set of demands, 150-200-200-400,[8] and the model outperformed the greedy algorithm (47.8% to 45.6%). Clearly this is due to the requirements pattern that made it important to plan ahead for the 400-response requirement of the fourth study. The 150-200-200-400 scenario was realistic given the demands the sponsor of our research had experienced. To gain further insight, we assumed a 1,000-1,000-1,000-1,324 scenario. Interestingly, the greedy algorithm generated feasible solutions in only 313 of the 1,000 replications. That is, by the time it got to the fourth study, it could not generate 400 respondents even if it solicited all available panelists. The model generated feasible solutions in 997 instances. In the 313 instances where both algorithms generated feasible solutions, the model outperformed the greedy algorithm (30.7% to 28.3%).

The above suggests (1) Monte Carlo simulation accurately predicted the field test; (2) the gains in the field test were due to the predictive model, not to the forward-looking optimization; (3) a greedy algorithm can work at least as well as the model when demands are constant; (4) the model works better when demands are not constant; and (5) the model can generate feasible solutions in cases of high demand requirements where the greedy algorithm does not.

7. Summary

We have developed an optimal contact model for maximizing online panel response rates and demonstrated that the model works in a real-world field test. In particular, it outperformed two benchmarks: random selection of respondents and a well-grounded heuristic. It generated this higher response without changing the sample characteristics of each study or the overall substantive responses.

Our work has important implications for researchers. First is that simple decision tree predictive models work well enough in application to improve performance. Second is that the rolling horizon approach is viable in a real-world marketing setting. Third is that simplifying an infinite horizon dynamic program to a finite horizon linear program produces strong results, i.e., there is "low hanging fruit" that can be captured by a simple model that captures essential phenomena. For practitioners, the most important implication is that response rates for online panels can be significantly improved through optimal contact models. The method is not black box. Both the predictive model (Table 5) and the optimal plan (Table 6) make intuitive sense.

Monte Carlo simulations suggest that the superiority of our model in the field test stemmed from the power of the predictive model, not to its forward-looking abilities. A myopic optimization heuristic, the greedy algorithm, would have done slightly better than our model in the field test. However, the field test, with a pattern of constant respondent demands from study to study, did not tax the forward-looking aspects of our model. Simulations of nonconstant demand patterns suggest that, in these cases, the model would outperform the greedy algorithm. The general lesson is that simple myopic algorithms work fine when the demands of the application do not require forward-looking optimization. However, the model advanced in this paper works well both in situations that require or do not require forward-looking optimization.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

The authors are grateful for invaluable research support provided by Sean Rhea, James Robbins, and Paul Wolfson, and for financial support from the UCR Sloan Center for Internet Retailing. They also thank Department Editor Jagmohan Raju, along with the associate editor and two reviewers for their helpful comments.

References

Ansari, A., C. F. Mela. 2003. E-customization. *J. Marketing Res.* **40**(2) 131–145.

Baker, K. R. 1977. An experimental study of the effectiveness of rolling schedules in production planning. *Decision Sci.* **8**(1) 19–27.

Bitran, G. R., S. V. Mondschein. 1996. Mailing decisions in the catalog sales industry. *Management Sci.* **42**(9) 1364–1381.

Blattberg, R. C., B. D. Kim, S. A. Neslin. 2008. *Database Marketing: Analyzing and Managing Customers*. Springer, New York.

Carleton, J. 2006. Panel integrity in the world of online market research. White paper, Frost & Sullivan, Palo Alto, CA.

Chand, S., V. N. Hsu, S. Sethi. 2002. Forecast, solution and rolling horizons in operations management problems: A classified bibliography. *Manufacturing Service Oper. Management* **4**(1) 25–43.

Davis, H. 2003. Are you talking to the right people? The importance of online panel quality. Quirk's Marketing Research Review (July/August) Article 1134, http://www.quirks.com/.

Elsner, R. 2002. *Optimiertes Direkt- und Datenbankmarketing unter Einsatz von mehrstufigen dynamischen Modellen* (*Optimized Direct and Database Marketing Using Multilevel and Dynamic Models*). Wiesbaden, Gabler, Germany.

Elsner, R., M. Krafft, A. Huchzermeier. 2003. Optimizing Rhenania's mail-order business through dynamic multilevel modeling (DMLM). *Interfaces* **33**(1) 50–66.

[7] This and other comparisons in this section are statistically significant, given 1,000 replications in the simulation.

[8] This was just simulated for the female panelists for simplicity.

Federal Supplement. 2002. *Albert v. Warner-Lambert Co.*, 234, F. Suppl. 2d 101.

Greenfield. 2005. Greenfield/Ciao online research barometer. http://greenfield.com/assets/pdfs/Barometer%2011-05.pdf.

Harris Interactive. 2006. An interview with our panel expert: The top questions about Harris interactive's internet research panel. http://www.harrisinteractive.com/partner/pubs/HI_Panel_Overview.pdf.

Hogg, A. 2003. Web efforts energize customer research. *Electric Perspectives* **28**(5) 81–83.

Lehmann, D. R., S. Gupta, J. H. Steckel. 1998. *Marketing Research*. Addison-Wesley, Reading, MA.

McClain, J. O., L. Thomas. 1977. Horizon effects in aggregate production planning with seasonal demand. *Management Sci.* **23**(7) 728–736.

Neslin, S. A., S. Gupta, W. Kamakura, J. Lu, C. Mason. 2006. Defection detection: Improving predictive accuracy of customer churn models. *J. Marketing Res.* **43**(2) 204–211.

Simester, D., P. Sun, J. Tsitsiklis. 2006. Dynamic catalog mailing policies. *Management Sci.* **52**(5) 683–696.

Stenbjerre, M., J. N. Laugesen. 2005. Conducting representative online research: A summary of five years of learnings. ESOMAR Worldwide Panel Research Conference, Budapest, April 17–19, ESOMAR, Amsterdam.

Sudman, S., B. Wansink. 2002. *Consumer Panels*, 2nd ed. American Marketing Association, Chicago.

Thornton, R. 2005. Online panels: The new frontier of B2B research. Ciao GmbH BIG Annual Conference, May, Ciao GmbH, Munich, Germany.