

Adept Adaptations:

Developing alternative strategies in adaptive survey designs

Jaki McCarthy, Tyler Wilson, and Andrew Dau

USDA's National Agricultural Statistics Service,
1400 Independence Avenue SW, Washington DC 20250

ABSTRACT

Many are turning to adaptive or responsive survey designs in their data collection processes to either minimize costs, reduce non-response bias or improve data quality. By identifying key design features affecting survey costs and errors, program managers can either proactively adapt or dynamically respond by altering those features in the survey data collection. However, for some surveys it may not be easy to identify alternative procedures to be employed in data collection. NASS has been testing the use of adaptive design in the Agricultural Resource Management Survey (ARMS). ARMS is an establishment survey with rich frame data that has been used to develop nonresponse propensity models. These models generate sampled records likelihood of response before data collection begins. Using these models a subset of records was identified and targeted with alternative data collection procedures.

ARMS, a long and challenging survey uses a mailout questionnaire with in person follow up for all initial nonrespondents. Because this survey already features in person interviews over a relatively long field period, it was not obvious what alternative data collection procedures we could use in an adaptive design. New alternative procedures included removing the initial mail contact, requiring the initial recruitment contact to be made in person and tailoring the recruitment to the specific sample unit. The procedures were based on feedback from field enumerators, input from survey program managers and principles of social psychology. Response rates and data quality indicators were compared between these groups. Response was not significantly increased in the treatment group most likely due to the limited number of contacts allowed. Preliminary review of the data quality showed some increases in the quality of reported data in the treatment group. Procedural changes are planned for the future, and may help other survey organizations develop effective alternative data collection strategies in adaptive designs.

INTRODUCTION

Many surveys are turning to adaptive or responsive survey designs in their data collection to either minimize costs, reduce non-response bias or improve data quality. By identifying key design features affecting survey costs and errors, survey managers can either proactively adapt or dynamically respond by altering those features in the survey data collection process. Thus multiple "phases" of a survey with differing conditions are combined in estimation and referred to as adaptive survey design (Groves and Heeringa, 2006). For example, in telephone surveys, response rates for certain subsets of records can be monitored during the field period. These groups may be defined in many ways but may be based on auxiliary data known about the sample cases. Groups with different response rates can be prioritized for differential efforts such as increased or capped number of call attempts, switches to different modes, or differential use of incentives (Mohl and LaFlamme, 2007, Wagner, et al, 2012, Lundquist and Samdal, 2013). However, for some surveys it may not be easy to identify alternative procedures to employ in data collection.

NASS has been testing the use of adaptive design in the Agricultural Resource Management Survey (ARMS). The ARMS is an annual survey which produces information about US agricultural resource uses, costs and farm financial conditions. Data from ARMS is used by many constituencies to evaluate the financial performance of the US farm sector. It informs agricultural policy decisions and provides the necessary background information to support evaluations of the relationships between agricultural production, resources and the environment.

ARMS is a multiphase survey, beginning with an initial screening phase, followed by a second phase collecting information on production practices. In the third phase farm operators provide data on farm operating expenditures, capital improvements, assets, and debt for agricultural production. In addition, operators provide data on farm-related income, government payments, the source and amount of off-farm income, and characteristics of themselves and their household. The focus of this paper is this third phase of data collection which will be referred to as ARMS.

ARMS annually samples over 30,000 farm operations. This long and challenging survey uses a mailout self-administered questionnaire with an informational cover letter followed by in person data collection for all mail nonrespondents. Interviews can last an hour or more. Because of the nature of the data collected, including both its sensitivity and burden to report, response rates for this survey have historically been lower than other NASS surveys. Therefore, NASS is interested in how best to allocate resources to improve response.

In order to implement an adaptive design, two elements are necessary. The first is identifying a set of cases to which alternative data collection procedures can be applied, and the second is developing procedures that will be applied. Because NASS contacts many of the same farms and ranches over time (both in ARMS and other surveys) rich frame data is available. In particular, NASS conducts the Census of Agriculture (COA) every five years. The COA includes all known or potential farms, and is one of the few NASS data collections that is mandatory. Unlike the population census, the COA is quite extensive and collects detailed information about the operation's production and characteristics. Information about the size, location, type or commodities, expenses, operator characteristics and more is collected. NASS has used COA data to develop predictive models which are used to generate a nonresponse propensity for each ARMS sample unit before data collection begins. (Earp, et al, 2014) Earlier exploration of adaptive design in ARMS began by focusing on the cases predicted to be most likely nonrespondents, defined as those with a predicted nonresponse rate of 70% or higher. However, efforts with these cases produced little return (Mitchell et al, 2015a, 2015b). Therefore efforts were redirected to cases that were still predicted to be nonrespondents but not the most difficult cases. These sample units with a nonresponse propensity between 50 and 70% were targeted for alternative procedures.

Because this survey already features an in person interview over a relatively long field period, it was not obvious what alternative data collection procedures we could use in an adaptive design. In order to implement an adaptive design we started with a review of the operational data collection procedures, and then used principles from social psychology to design alternative procedures.

DEVELOPMENT OF ALTERNATIVE DATA COLLECTION PROCEDURES

Current ARMS procedures are designed with consideration of data quality, response and costs. Obviously, data that can be collected by mail with a self-administered form is far less costly than data collected using interviewers, either by phone or in person. So NASS initially mails out a survey package to the entire sample. The package consists of a paper form with a standard cover letter requesting participation, and emphasizing the importance of the data. Mail response rates can be substantial, and over a third of the responses have historically been collected through the mail. After an initial period to allow mail responses, any mail nonrespondents are assigned to field interviewers for follow up. Interviewers are instructed to visit or telephone to set up appointments for interviews at their discretion. They will then typically call to schedule an appointment and conduct in person interviews to collect the data. Thus NASS has instituted overall procedures applied to all sample units, to minimize costs while still including in person interviews to help increase response rates.

However, when viewed from the perspective of respondents, the procedures may look quite different. The first contact a potential respondent has for this survey is an unsolicited mailing. The mail package contains the paper form, typically 24 pages or longer with a cover letter that describes the importance of the survey to the overall agricultural sector. The only personalization of the mailing is in the addressee name and address and salutation. Respondents are requested to complete and mail back the survey. ARMS is typically not a mandatory survey (although in the year of this experiment it was combined with another survey and carried mandatory reporting authority).

At this point, the respondent can either complete the form and return it, set the form aside, or decide not to respond. For any sampled unit that does not return a form, a second mailing is delivered. Again, at this point, a respondent who has set the form aside, may decide to complete and return it, or may again refuse to complete it. After this second mailing, respondents who still have not returned the form are contacted by an interviewer and asked to schedule an interview to complete the form.

Following prior year data collections, interviewers complained that some respondents had already decided to refuse participation before being contacted by interviewers for data collection interviews. This made it more difficult to obtain participation than if their initial contact was from an interviewer as illustrated by this quote from field office staff: *“For many people, if they knew what they were getting into they wouldn’t do it. Mailing it ahead, they know what you’re coming to ask, so they refuse. Others want the extra time to get started. But overall, consensus is that mailing ahead hurts their ability to conduct a field interview.”* Viewed from the respondents’ perspective, when the interviewer calls to schedule a follow up interview appointment, they may have already received the questionnaire and decided to refuse twice.

Interviewers also noted that some respondents asked whether they could complete the form on their own without the interviewer. Interviewers are trained to conduct follow up interviews and are not given the option for alternative methods of data collection.

In order to design alternative data collection procedures to use in adaptive design, we took the perspective of the respondents to alter how the survey recruitment process appears to them. Our new procedures sought to incorporate principles of social psychology to increase respondents’ motivation to respond.

SOCIAL PSYCHOLOGY AND MOTIVATION TO RESPOND

The subset of the sample that was predicted to have between 50-70% nonresponse propensities was identified using the nonresponse propensity models. These operations are not the most difficult to obtain response from, but are still more likely to be nonrespondents than not. Half of this sample received the standard data collection procedures. The second half of these targeted operations used the following alternative data collection procedures:

- No initial questionnaire mailing,
- Cases identified for interviewers,
- First contact was by interviewer visiting in person, up to two attempts allowed,
- Interviewers asked to tailor introduction and recruitment to individual operation,
- Interviewers provided token incentives to provide during introductions,
- Respondents given option to drop off and pick up questionnaire or schedule an interview for data collection,
- If respondents refused, interviewers were required to leave questionnaire and return at a later date to pick up blank (or completed) questionnaire.

Groves, Cialdini and Couper (1992) have discussed how principles identified by social psychologists (liking, reciprocity, authority, social validation, consistency, etc.) can impact a respondent’s willingness to comply with requests for survey participation. Our new procedures were designed to leverage these principles in the survey recruitment process to increase the likelihood of survey response. For example,

- Liking: people are more likely to comply with requests from people that they like, rather than people they do not like or people they do not know,
- Authority: people are more likely to comply with requests from sources of authority,
- Consistency: people are prone to act in a manner they perceive to be consistent with their other actions or viewpoints,
- Reciprocity: people will feel a greater sense of obligation to someone who has already done something for them.

We hypothesized that response would be increased if we could increase the “liking” that the respondent had for the interviewer. In order to increase the rapport between the interviewer and respondent, for the targeted subset of respondents, they were not included in the initial questionnaire mail out. Instead, interviewers were required to make the initial contact in person rather than by telephone or mail. In contrast to the generic survey cover letter, they were also asked to tailor the recruitment to the individual respondent, using examples of uses of the data that were relevant either to operators in that area or to that type of operation. As in their usual practice, they were also to cite the USDA and NASS as the sponsor of the survey, in order to emphasize the authority of the request.

The consistency principle was relevant in that we did not want respondents to form negative opinions of the survey and their participation before being contacted by the interviewer. So for these predicted nonrespondents, we did not mail them a questionnaire to give them the opportunity to decide against participation on their own. We did not want potential respondents to think that complying with the interviewers’ request was inconsistent with their existing views.



Figure 1. Treatment Data Collection Materials for Interviewers

In order to engender a sense of reciprocity, interviewers were provided with token incentives bearing the NASS logo (e.g. small notepads, ice scrapers, work gloves, rain gauges, etc.) which were to be provided up front in the initial contact. In addition, after making contact, they were allowed to either leave the form for the respondent to complete on their own or schedule an interview based on the respondent’s preferences. We also hypothesized that the interviewer’s effort in having traveled out to the sampled operations would also help promote the sense of reciprocity. Even those that refused the first contact would also have a second in person interviewer visit where they would have to reiterate their refusal to the interviewer who had again traveled out to them in person.

In addition to increasing response rates, alternative data collection procedures also have the potential to change the quality of the data reported. This could potentially either increase or decrease the quality of the data. If respondents

are reluctant, they may put less effort into reporting and produce lower quality data. However, if respondents' motivation to report has been increased, as we hoped with our procedures, then we might expect the effort put into reporting to be greater and the quality of the data increased.

METHODS

Nonresponse propensity scores were generated for all ARMS sample units based on classification tree models. This study included only sample units with predicted nonresponse propensities between 50-70%. Each unit was randomly assigned to either a treatment or control group. The control group received the standard data collection procedures, and the treatment group received the alternative procedures described above. In addition, for each treatment case, interviewers completed a form indicating whether they had been able to use the experimental procedures, when they had contacted the operation, whether they had successfully made contact with the respondent, and any other comments for that specific case.

Next, measures of data quality were compared between the treatment and control respondents. Target items were identified for this preliminary analysis. The items were operator seed expense, the value of owned land and the depreciation expense. These items were chosen because they encompass a variety of reporting and data processing characteristics. For each respondent the number of times the target items were imputed or edited was calculated. Those with higher imputation and edit rates were assumed to be of poorer quality.

RESULTS

Response Rates

There were 774 records assigned to our treatment group. Field office staff was asked to identify records for which the treatment procedures could not be used (for example, records for which special data handling was previously in place, records for which data collection was being conducted in conjunction with other survey contacts, etc.) This reduced the treatment sample to 659 records. Overall there was not a significant difference in the response rate between the treatment and control groups with the treatment group.

Table 1: Response rates for Treatment and Control Group

	Treatment (n=659)	Control (n=778)
Completion (1)	59%	58%
Refusal (2)	34%	33%
Inaccessible (3)	6%	6%
Out-of-Scope (4)	1%	2%

As predicted by our nonresponse propensity models, both groups did have a higher nonresponse rate than the part of the sample that was not included in our experiment. The remaining records that were not included in our experiment had a response rate of 75%.

However, only 57% of the treatment sample was contacted in person by an interviewer, (45% in the first attempt and an additional 12% in the second attempt). This suggests that allowing only 2 contact attempts for the initial recruitment visit was not sufficient. Of the group that was contacted, the response rate was 70%. For records where no contact was established the response rate was only 41%. Thus, if more cases had been contacted we might have been able to increase response rates further. Feedback collected from the field office staff also supported the idea that more contacts attempts are needed. In addition, field staff reported that interviewers thought eliminating the initial mailed questionnaire was beneficial. We also learned from the field feedback that some of the interviewers had misinterpreted the instructions and thought that they were required to drop off and leave the form with

respondents and could NOT conduct interviews. For those respondents that would have preferred interviews (either immediately or at a later scheduled time) this may have also adversely affected response rates.

Field office staff also suggested that additional contact attempts would have been useful. They also suggested allowing telephone contacts when samples had been attempted unsuccessfully in person several times.

Data Quality

In order to examine data quality, the edit and imputation rate was compared between our control and treatment respondents for three target variables. This was done by recording the number of times the data reported by a respondent was different from the data in the final dataset. The percent of the items changed is shown in the tables below for our control and treatment groups. (Note: the number of items depends on if this item is applicable for the specific respondent, i.e. livestock operations may not have seed expenses, not all operations will have owned land, etc.) Table 2 below shows that 2.63% of usable records in the treatment sample had changes to the reported data from seed expense. A higher percentage (5.03%) of records in the control group had changes to the reported data.

Table 2: Seed Expense

Mode	Treatment			Control		
	<i>N</i>	<i>Observations with Change</i>	<i>Percent</i>	<i>N</i>	<i>Observations with change</i>	<i>Percent</i>
Mail	64	2	3.13%	135	10	7.41%
Phone	7	0	0.00%	9	0	0.00%
Face	344	8	2.33%	256	10	3.91%
Web	4	1	25.00%	37	2	5.41%
Total	419	11	2.63%	437	22	5.03%

Table 3 shows that 21.96% of usable records in the treatment sample had changes to the reported data from the *Value of Owned Land*. Once again, a higher percentage (24.49%) of records in the control group had changes to the reported data.

Table 3: Value of Owned Land

Mode	Treatment			Control		
	<i>N</i>	<i>Observation with Change</i>	<i>Percent</i>	<i>N</i>	<i>Observations with Change</i>	<i>Percent</i>
Mail	64	18	28.13%	135	45	33.33%
Phone	7	3	42.86%	9	4	44.44%
Face	344	70	20.35%	256	50	19.53%
Web	4	1	25.00%	37	8	21.62%
Total	419	92	21.96%	437	107	24.49%

Lastly, the third variable analyzed was *Operator Depreciation*. *Table 4* shows that 32.22% of records in the treatment sample had changes made from the reported data. A similar proportion (32.27%) of reported responses in the control group were changed. *Operator Depreciation* appears to be a very difficult question for respondents to answer regardless of the data collection strategy. This is because depreciation is often a number calculated only for tax filing purposes and generally not a number tracked during the year. However, a variable such as this reveals that data quality is often a function of the type of question asked and not how the answer was collected.

Table 4: Operator Depreciation

Mode	Treatment			Control		
	<i>N</i>	<i>Observations with Change</i>	<i>Percent</i>	<i>N</i>	<i>Observations with Change</i>	<i>Percent</i>
Mail	64	18	28.13%	135	39	28.89%
Phone	7	1	14.29%	9	6	66.67%
Face	344	114	33.14%	256	92	35.94%
Web	4	2	50.00%	37	4	10.81%
Total	419	135	32.22%	437	141	32.27%

Although these indicators provide insight into how different types of questions are functioning in regards to hard-to-respondents, a general test of significance (Wilcoxon, 0.05 Level) reveals that any differences between samples were not significant. *Table 5* illustrates the p-values after testing the rate of changes between samples.

Table 5: Wilcoxon Two-Sample Test Two Sided P values

Variable	Normal Approx.	T Approx.
Seed Expense	.1752	.1847
Value of Land	.5785	.5785
Depreciation	.6503	.6506

If time and resources allowed, it would be insightful to test other measures of data quality, and perhaps comprehensively test all rates of change in every (over 1400) item code response in ARMS.

These tables show first that the percent of changes is different for each item, with far fewer edits made to the operator seed expense than either of the other two items. Comparing our treatment and control groups, the direction of the data quality indicators is in the right direction (i.e. fewer changes) for our treatment respondents for operator seed expense and value of owned land but not significant. This demonstrates that quality varies for different items

but that our treatment procedures have potential to improve quality for some items. It is possible that interviewers can help respondents to understand what is being asked or encourage them to report for difficult items in our alternative procedures. However, some items, like depreciation expenses are difficult regardless of the way they are collected.

DISCUSSION

Our experiment with using alternative data collection procedures on a targeted subset of the ARMS sample yielded disappointing results. However, a number of ways to improve the procedures were identified and another experiment is underway in the following year of ARMS. Improvements in the process to allow additional contact attempts and stress that respondents have the option of either completing the form themselves or in an interview will be instituted. It is also crucial to track how experimental procedures are actually carried out in data collection. In addition, ARMS was conducted in conjunction with another survey in the year of our experiment and had a mandatory reporting authority. This will not typically be the case, and in future years, ARMS will continue to be collected on a voluntary basis. The mandatory reporting authority may have increased the expected response rate for our control group and masked some of the positive effect of our treatment procedures.

As adaptive and responsive survey design are increasingly introduced in survey data collection it will be important to continue work in several complementary areas. Methods to identify the goals and appropriate subgroups of the sample to target will be necessary. In our case we identified those records which were likely to be nonrespondents but perhaps sensitive to alternative procedures. We hoped to increase the response rate for this group, with the ultimate goal being reduced nonresponse error. However, we also examined data quality relative to our procedures. If increasing data quality is the goal, we may need to rethink the way we target sample members for alternative data collection procedures. There are numerous ways to measure data quality, we have only identified one in our preliminary analysis. Identifying the records most likely to introduce error into survey estimates may be an alternative approach to applying adaptive survey design.

Along with identifying target sample subgroups, more effective data collection procedures must be developed for these records. This is often a difficult task, as data collection is often already planned to optimize existing resources. However, as outlined in this paper, we have attempted in our experiment to view the data collection recruitment from the respondents' perspective. That is, developing procedures that are not just more resource intensive, but build on principles of social psychology. This view of the recruitment as an interaction between the respondent and the survey organization and its interviewers can help shape the initial survey recruitment. And this may be a more effective approach than viewing the survey solely from the viewpoint of survey managers and their overall budgets. More customized approaches to the targeted records may be necessary.

REFERENCES

- Earp, M., Mitchell, M., McCarthy, J., and Kreuter, F. (2014) Modeling Nonresponse in Establishment Surveys: Using an Ensemble Tree Model to Create Nonresponse Propensity Scores and Detect Potential Bias in an Agricultural Survey. *Journal of Official Statistics*, Vol 30(4), 701-719.
- Groves, R., Cialdini, R. and Couper, M. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, Vol 56, 475-495.
- Groves, R. and S. Heeringa. (2006). "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: 439-457.
- Lundquist, P. and Sarndal, C. (2013) Aspects of Responsive Design with Applications to the Swedish Living Conditions Survey. *Journal of Official Statistics*, Volume 29(4), 557-582.

Mitchell, M., Ott, K., Ridolfo, H. and McCarthy, J. (2015a). Targeted Data Collection Efforts for the 2013 ARMS III, RDD Research Report RDD15-02, USDA, National Agricultural Statistics Service, Washington, DC.

Mitchell, M., Ott, K. and McCarthy, J. (2015b). Targeted Data Collection Efforts for the 2012 ARMS III, RDD Research Report RDD15-01, USDA, National Agricultural Statistics Service, Washington, DC.

Mohl, C. and F. Laflamme. (2007). Research and Responsive Design Options for Survey Data Collection at Statistics Canada. In Proceedings of the Joint Statistical Meetings: American Statistical Association. Alexandria, VA: American Statistical Association.

Wagner, J., West, B.T., Kirgis, N., Lepkowski, J.M., Axinn, W.G., and Kruger-Ndiaye, S. (2012). Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection. *Journal of Official Statistics*, 28(4), 477-499.