

They Spoke, We Listened: Reducing Respondent Burden Using Previously Reported Data

Emilola J. Abayomi¹, Steve Maliszewski¹, Lindsay Kreiner¹, Todd Ballard¹

¹United States Department of Agriculture-National Agricultural Statistics Services, 1400 Independence Ave. SW, Washington, DC 20254

Abstract

The National Agricultural Statistics Service (NASS) administers the Crops Acreage, Production and Stocks (APS) survey for the purpose of providing estimates for crop acreage, yield, and production. Respondents are asked questions about acres operated, more specifically land owned, land rented from others, land rented to others, total acres operated and cropland, each quarter. Redundant questions such as acreage have increased respondent burden. This research was conducted to determine whether respondent burden can be reduced by capturing the data in June (base month) and then reusing these data throughout the survey cycle for September, December and March. This study will explore the utilization and validity of previously reported data (PRD) and the effects on respondent burden. The results of this research and future direction to address respondent burden for the Agricultural Survey will be discussed.

Key Words: previously reported data, respondent burden, web-based survey, agricultural survey

1. Introduction

The National Agricultural Statistics Service (NASS), along with other survey-driven organizations, has seen a steady decline in response rates over the years. Various factors have led to this decline. NASS has taken proactive measures to address one of those factors – increased respondent burden. Respondent burden, defined as the time, effort and cost required for respondents to complete a survey (Bradburn, 1978), not only impacts response rates but also data quantity, quality and survey sustainability.

In an effort to alleviate respondent burden, previously reported data (PRD) were one of four data sources identified by NASS that could be used instead of collecting data from respondents or could be used as an aid during reporting. PRD are data collected from a respondent during an earlier survey period. It is not a new phenomenon to the survey community and have been utilized in a variety of ways from pre-printing on self-administered paper and web questionnaires (Holmberg, 2004 and Mooney et al., 2009) to telling the respondent the previous survey data directly before collecting a current response. As a result of several research experiments, NASS has operationalized PRD for edit checks, pre-response prompts, and in lieu of asking questions (Gottschall, 2009).

Research results have been mixed on PRD administration for reducing burden, data collection efficiency and improving data quality however the consensus is PRD can be beneficial if used in a non-biasing manner. This paper will focus on efforts taken to alleviate respondent burden using previously reported data on the web instrument. The validity of PRD and the respondents' response process when presented with PRD are evaluated.

2. Methods

The Crops Acreage, Production and Stocks (APS) survey is a quarterly survey administered in June – the base survey, September, December and March. This survey collects data on a variety of small grains and row crops, so it targets operations with cropland and storage capacity. It is administered in all states except Hawaii. The decision to use this survey was based on feedback from respondents about the acres-operated questions consisting of land owned, land rented from others, land rented to others and cropland. These questions are asked throughout the survey cycle; however, respondents have expressed that their reported acres rarely change between quarters.

The efforts for this study concentrated on Missouri, Illinois, Iowa, Wisconsin and Minnesota, using the Computer Assisted Web Interview (CAWI). For the pilot study starting June 2018, all respondents received a blank online questionnaire and were responsible for providing data for the acres-operated section as seen in figure 1. For September 2018 and the follow-up surveys, the data from the previous survey will be carried forward for this section, and respondents will have the opportunity to update or verify the information.

Figure 1: Acres Operated Questions for the Online Crops APS survey

The screenshot displays the 'Acres Operated' section of the 'AGRICULTURAL SURVEY - JUNE'. The interface features a navigation menu on the left with sections: Introduction, Acres Operated (selected), Crops, Stocks, Value of Sales, Change in Operator, and Conclusion. The main content area contains the following text and questions:

Please report total acres operated under this land arrangement.

- On June 1, how many acres did this operation:
 - Own? acres
 - Rent or Lease from others or use Rent Free? EXCLUDE land used on an animal unit month (AUM) basis, BLM and Forest Service land. acres
 - Rent to others? acres
- Then the total acres operated on June 1 was: acres
 - Does this include the farmstead, all cropland, woodland, pasture land, wasteland, and government program land? Yes No

The bottom portion of the screenshot shows the same 'Acres Operated' section with question 4:

The remaining questions in this survey refer to the total acres operated.

- Of the total acres operated, how many acres are considered cropland, including land in hay, summer fallow, cropland idle, cropland used for pasture and cropland in government programs (INCLUDE land in the Conservation Reserve Program (CRP) and in the Wetland Reserve Program (WRP))? acres

Two methods were used to justify pre-filling responses for subsequent surveys. First, four quarters of data for June 2017, September 2017, December 2017 and March 2018 were pulled for quantitative analysis. During the pilot study, respondents will receive PRD from the immediate prior survey to ensure they are receiving the most up-to-date information. For this reason, the quarters were paired for analysis: June 2017 to September 2017, September 2017 to December 2017, and December 2017 to March 2018. Only useable records were retained. Usable records were categorized as (1) complete during the prior survey period, (2) complete by the same operations in both months of comparison and (3) complete by an operator, spouse, partner or bookkeeper/accountant.

The variable cropland acres was evaluated by using descriptive statistics such as the number of usable records between quarters and percentage of records reporting change. A paired t-test for each month's pair was applied to determine whether a significant difference in the means for cropland acres was detected between quarters. States were grouped for analysis to reflect the defined regions established by NASS. The Heartland region consists of Missouri and Illinois and the Upper Midwest region consists of Iowa, Minnesota and Wisconsin.

Next, usability interviews was conducted using the web instrument. Usability testing is a qualitative research procedure used to determine how well a product mirrors a user's expectations by accessing intuitiveness, efficiency and satisfaction. Moderators were responsible for recruiting users from the NASS list frame. Users were informed of the intent of the testing during the recruitment process. Participation was voluntary. Six usability interviews were conducted in Iowa, Illinois and Missouri, all of which were conducted in person. Prior knowledge of the web instrument was not necessary for participation; however, internet access was a requirement.

Moderators were instructed to give the login information to the users and have users complete the survey as they normally would under ordinary conditions. During the interviews, users were asked to complete various tasks while the moderator asked concurrent and retrospective probes to determine whether features of the web questionnaire were useful and intuitive. The moderator also observed and noted the users' reactions as they worked through the survey. Each interview took approximately 30 minutes to conduct. The users' responses, reactions and feedback were compiled and analyzed following the interviews. The goal was to discover common patterns in the data that uncovered problematic features and develop strategies to correct those problems accordingly.

3. Results

3.1 Quantitative Analysis Results

Tables 1-4 detail the comparisons of PRD for a given quarter compared to reported data for a subsequent month in the Heartland and Upper Midwest regions.

Quarters	Usable Records	Percentage of Records Reporting Change	Difference in the Means	Absolute Difference in the Means
June 2017 to September 2017	1071	50.89%	10.04	289.26
September 2017 to December 2017	2155	77.26%	6.33	203.98
December 2017 to March 2018	1997	73.66%	4.98	219.14

As seen in table 1, more than 50% of operations changed the value of cropland acres in all quarter comparisons. The difference in the means ranged from 4.98 to 10.04 with the largest difference occurring from June to September. The absolute differences in the means were larger ranging from 203.98 acres to 289.26 acres. This indicates the reported data are widely scattered from period to period. A significance test was conducted to determine whether these differences were significant.

Pairs	n	Mean	Std Dev	Std Err	T Value	Pr > t
June 2017 to September 2017	1071	10.04	1330.80	40.66	0.25	0.8049
September 2017 to December 2017	2155	6.33	718.30	15.47	0.41	0.6824
December 2017 to March 2018	1997	4.98	760.20	17.01	0.29	0.7696

Table 2 summarizes the paired t-test for the matched records from period to period. Each row in the table displays results from a different test. Reported cropland acres in June were compared to reported cropland acres in September in row 1. Reported cropland acres were higher in September (1393.25) than June (1383.20), but the difference was not significant indicated $p = 0.8049$. Reported cropland acres in September were compared to reported cropland acres in December in row 2. Reported cropland acres were higher in December (1163.06) than September (1156.73), but the difference was not significant. Reported cropland acres in December were compared to reported cropland acres in March in row 3. Reported cropland acres were higher in March (1073.17) than December (1068.19), but the difference was not significant. These results indicate the reported acres were consistent between consecutive quarters.

Quarter	Usable Records	Percentage of Records Reporting Change	Difference in the Means	Absolute Difference in the Means
June 2017 to September 2017	1546	67.79%	-20.08	210.49
September 2017 to December 2017	2980	70.64%	5.07	139.91
December 2017 to March 2018	2629	71.05%	-41.98	179.91

As seen in table 3, more than 67% of operations changed the value of cropland acres in all quarter comparisons. The difference in the means ranged from -41.98 to 5.07 with the largest difference occurring from December to March. This is not a surprise because March marks a new crop year where rental agreements for land may change. The absolute differences in the means were larger ranging from 139.91 acres to 210.49 acres. This indicates the reported data are widely scattered from period to period. A significance test was conducted to determine if these differences were significant.

Quarter	n	Mean	Std Dev	Std Err	T Value	Pr > t
June 2017 to September 2017	1546	20.08	980.60	24.93	0.81	0.4208
September 2017 to December 2017	2980	5.06	674.10	12.34	0.41	0.6816
December 2017 to March 2018	2629	-41.98	642.40	12.52	-3.35	0.0008

Table 4 summarizes the paired t-test results for the Upper Midwest region. Reported cropland acres were lower in September (932.19) than June (943.27), but the difference was not significant with $p = 0.4208$. Reported cropland acres were higher in December (877.95) than September (872.89), but the difference was not significant. These results indicate the reported acres were consistent for these survey periods. Reported cropland acres were lower in March (908.94) than December (950.92) and the difference was significant indicated with $p = 0.0008$. This needs to be explored more, but it is hypothesised that the significant difference in means is due to rental agreement changes that occur in March.

Overall, the cropland variable appears to be a good candidate for pre-printing. Although operators tend to change this variable with a significant difference seen in the Upper Midwest region from December to March, we are confident to proceed with this study, because respondents will have an opportunity to update the information.

3.2 Usability Interview Results

The majority of users are the primary producers of the operations though two users share the decision-making with a spouse. It was assumed that respondents complete web surveys similar to telephone surveys. They do not prepare beforehand as the task tends to be spur of the moment and dependent on available time. This held true because one user completed the questionnaire using records, while the others completed the questionnaire from memory. Users were approximately the same age, and the majority were male as seen in Table 5 below.

Age	Male	Female
18-29	0	0
30-44	1	0
45-64	3	2
65+	0	0
Total	4	2

Exploratory questions were administered to understand the motivation for users to utilize the online survey. Top reasons for completing the questionnaire online were convenience and ease. As noted by a user, “I can think about and check information, which you can’t do on the phone and it’s easier than completing it by hand.” All users were familiar with online reporting, and all but one preferred web reporting to other modes of completion.

The first screens of the web instrument ask the user to update or verify information about the operation and partnerships, followed by screener questions for the total acres operated. The screen seen in Figure 2 starts the content of the survey.

Figure 2: Instruction Text



Instructions displayed in figure 2 were placed at the top of screens with PRD to inform users that PRD were present and to verify or update the information accordingly. Through observation, all but one user did not read the instructions. The one user that read the instructions read them aloud slowly. When probed about not reading the instructions, users commented that they were familiar with the instrument and questions and wanted to save time by bypassing the instructions. One user noted that she would have read the instructions if it was her first time completing the online survey while another user stated

he noticed the pre-filled responses first and jumped straight to the questions. Three users suggested bolding the instructions in order to draw the eyes up to the instructions.

Figure 3: Displayed Data and Suppressing Change

The screenshot shows the 'Acres Operated' section of the 'AGRICULTURAL SURVEY - JUNE'. The sidebar on the left lists sections: Introduction (checked), Acres Operated (selected), Crops (checked), Genetically Modified Seed Use (checked), Stocks, Value of Sales, Change in Operator, and Conclusion. The main content area contains the following text and form elements:

Please report total acres operated under this land arrangement.

Last quarter you reported the following acres. If the acres have changed then update with the current acres and click next. If the acres have not changed then click next.

1. On June 1, how many acres did this operation:

- a. Own? acres
- b. Rent or Lease from others or use Rent Free? Exclude land used on an animal unit month (AUM) basis, BLM and Forest Service land. acres
- c. Rent to others? acres

2. Then the total acres operated on June 1 was: acres

a. Does this include the farmstead, all cropland, woodland, pasture land, wasteland, and government program land?
 Yes No

Two major concerns of using PRD were of interest to the agency: data disclosure and failure to update the displayed information. First, the respondents' initial reaction of viewing the data seen in figure 3 was gauged. When probed, participants said they felt comfortable with having their information displayed in the instrument. They had no problem with recognizing the data as their own and were aware the values displayed were from a previous survey. One respondent commented that three people (primary producer, spouse and partner) tend to fill out the surveys for the operation and that the PRD would be helpful, because it would make all of them consistent in reporting if a different person filled out the survey in different quarters. Two respondents commented that they were comfortable with viewing their data online and did not have concerns with privacy. The second concern was the risk that respondents would not update the displayed information. However, two respondents changed their values. One was a hard change in which the respondent said the displayed information was wrong. One was a soft change in which the respondent initially changed the value, consulted records and then changed the value back to the original. The respondents that did not change the PRD said the information had not changed since the last survey but would change the information in the future if necessary.

Figure 4: Users' expectations

The crops screen shown in figure 4 is displayed after the acres operated section – the only section prefilled with responses. In this section, operators are to report information such as acres harvested and to be harvested, acres planted and to be planted, etc. for various crops. Some respondents were expecting to see pre-filled responses for the entire survey. They were flustered as to why the entire survey was not filled out. This should be considered as the agency moves forward with incorporating PRD into surveys.

4. Conclusion

No gold standards for utilizing previously reported data exist. This pilot study will serve as the basis for NASS to set limits on how and when to utilize PRD specifically for web surveys. Crops APS appears to be a good candidate for PRD usage as seen by the results from the quantitative analysis. Results from usability interviews also confirm that PRD appeared to decrease respondents' perceived burden when used as an aid during reporting. Debriefing questions were asked immediately after the usability interviews and all respondents exclaimed that PRD were helpful and they would like to have it, available, for Crops APS and other NASS surveys. Comments included but was not limited to "It makes things easier" and "Saves time." These findings suggest that PRD can be used as an incentive to complete the online survey and could possibly reduce sample attrition and stabilize or boost response rates over time. However, continued PRD use should be monitored to determine long-term effects on published estimates. This analysis provided insight into comparability of PRD data for a survey cycle but it was limited to specific regions and variables. In the future, this study will include all states to determine how published estimates may or may not be affected throughout the United States.

References

- Bradburn, N. 1978. Respondent Burden. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 1978: 35-40.
- Gottschall, C. 2009. Previously Reported Data Usage in NASS Field Offices. Census and Survey Division, Research Report – DCB-09-01. United States Department of Agriculture, National Agricultural Statistics Service.

Holmberg, Anders. 2004. Preprinting Effects in Official Statistics: An Experimental Study. *Journal of Official Statistics*. 20(2): 341-355

Mooney, Geraldine M., Melissa Krakowiecki, and Deborah Trunzo. 2009. Does Prefilling Responses on a Longitudinal Establishment Survey Stem Sample Attrition? *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 2009:5917-5924.