Developing and Evaluating a Short Form: Results and Recommendations From Tests of a Form Designed to Reduce Questionnaire Length

Danna Moore, Kathy Ott, and Arina Gertseva[1]

Abstract

Every five years, National Agricultural Statistical Service (NASS) conducts the Census of Agriculture (COA). The questionnaire is quite lengthy and detailed, although any given respondent should not have to complete all sections of the form. For the 2017 COA, NASS endeavors to create a shortened form with the goal of reducing the perception of burden and tailoring the form to specific operations. This paper discusses approaches taken to reduce the size of the form, testing and evaluation, and testing outcomes. Approaches included reducing length by combining historically separate sections of the form, removing sections of the form, and then using frame data to reduce the population eligible to receive it, removing pre-printed items, and requiring more written responses in tables. Questionnaires were tested by random assignment of agricultural operations to form experimental treatments. Results from evaluations of these strategies are presented. Findings demonstrate a preferred format for reduction and design guided by data quality outcomes.

Key Words: Questionnaire testing, data quality, item nonresponse

1. Purpose and Research Objectives

A long held conjecture is reducing questionnaire length reduces response burden as respondents answer fewer questions thus spending less time on the survey task. There are previous attempts by authors to estimate response based on number of pages in a questionnaire (Bruvold and Comer, 1988). Previous studies show the effects of reducing questionnaire length on study measures are unpredictable. Generally, research shows increased length and increased complexity are factors associated with increased respondent burden questionnaires that tend to reduce response (Axhausen and Weis 2010; Fox et al, 1988; Heberlein and Bumgartner, 1978). Questionnaire length has been described as one of the largest costs of being a respondent (Dillman et al (2014) and questionnaire complexity in USDA agriculture surveys contributes to both unit nonresponse and item non-response (NRC, 2007). Multi-purpose agricultural surveys such as the COA and Agricultural Resource Management Survey (ARMS) that collect agricultural production, farm operation and financials, and farm household information are comprised of questions that may be technical, hard to retrieve, hard to estimate or may be unclear which greatly adds to the burden respondents perceive or experience. Different lengths of survey forms may produce differential unit response and may also produce non-equivalent item responses. Questionnaire visual design elements in the form of commodity code lists,

[1] Danna L. Moore and Arina Gertseva, Social & Economics Sciences Research Center, PO Box 1801, Washington State University, Pullman, WA 99164-4014, mooredwsu.edu; Kathy Ott, NASS, USDA, 1400 Independence Ave, SW, Washington, DC 20250-2001Kathy.Ott@nass.usda.gov.

matrix reporting tables and placement of instructional information needed by respondents to fully answer survey questions may influence item response quality and overall data quality. Leverage saliency theory (Groves et al. 2004) postulates survey topic saliency and interest plays an important role in response, however, for agricultural surveys, response is additional influenced by extraordinary response burden which may be offsetting saliency as agricultural operations are asked to be historical (repeat) respondents as the largest producers of specific commodities are critical to agricultural measures. It is highly recommended that researchers and survey designers who rely on questionnaire redesign strategies be aware of the effects of questionnaire changes, and where possible test, understand, and address the impacts. Towards this end, this paper demonstrates the impact of specific questionnaire changes to shorten and reduce burden of complex farm business questionnaire forms.

In this study we also concentrate on changes to questionnaires as an on-going process that can have impacts on survey data quality and data accuracy. Biemer and Lyberg (2003) describe accuracy as a premier feature of data quality that is rarely measured, is difficult and expensive to measure, and that to assure quality is reliant on quality processes. To achieve continuous improvement in survey operations requires feedback and a learning process. In overviewing data collection methods and survey quality, deLeeuw and Collins (2012) and Snikers, et. L. (2013) provide a number of factors that affect the usefulness of a survey, which includes those we intend to focus our attention to: the method of data collection, questionnaire imperfections, processing errors and errors in interpretation. To make further strides in the reduction of survey burden and sustain or improve survey quality, agricultural respondents are considered a special population which requires hand crafted questionnaires. Pretesting and experimental tests are used to evaluate the directional impact of questionnaire changes. The Census of Agriculture long form is 24 pages and is not part of our study, however it is the original source form. This paper reports the effects of an experiment of a Census of Agriculture 16-page condensed form tested in comparison to a shorter 12-page reformatted form for self-reported agricultural operation information in a nationwide sample of farm and ranch operations.

In many survey organizations, and NASS in particular, there has been a movement to look for causes of error, control for error, and make improvements by studying survey error and quality. The main research question in our study was “Did the reformatted shorter 12-page form improve reporting by respondents?” To further evaluate the performance of the revised forms, we considered numerous factors: 1) unit response rates across the two forms; 2) item data quality in terms of nonresponse on each question across the form; 3) operational items needed for consolidation and data cleaning (i.e. use of commodity codes, blanks, multiple marks, non-responsive answers); 4) the number of questions answered; 5) the number of questions where the numeric information was correctly reported; and 6) the performance on screening questions. To our knowledge this study is unique as it additionally tests and reports item response associated with alternative presentations of commodity code tables in agricultural questionnaires. While this last feature is used extensively in agricultural surveys, it is also a feature common and needed in other non-agricultural enterprise surveys. Similarly, there are few published studies reporting reduced questionnaire length effects on detailed agricultural categories and numeric commodity measures. Two sets of hypotheses are assessed. The first evaluates the impact of reducing questionnaire length and the second evaluates the placement of commodity code reference

information influencing item level responses. This study tests experimentally: 1) Hypothesis test 1, reduction of questionnaire length. The null hypothesis, H_0: Shorter questionnaire length will increase unit response. The alternative, H_A: Shorter questionnaire will show no difference in unit response; 2) Hypothesis test 2, reformatting of commodity code reference list. The null hypothesis, H_0: Placement of specific commodity code list at question reduces item nonresponse. The alternative hypothesis, H_A: Placement of specific commodity codes at question shows no difference.

2. Background

The Census of Agriculture (COA) is the only source of statistics for US Agriculture that provides information on production agriculture at the county level in every state and US territory (C-FARE, 2007). Every 5 years, National Agricultural Statistics Service (NASS) conducts the COA with farms of all sizes across the entire US to collect data to meet legal mandates, policy and program needs, and demands of a wide breadth of users. The COA is designed to provide a complete and comprehensive picture of American agriculture. To continue to provide data with high utility and to satisfy end users, it is requisite for NASS to adapt and change the COA to capture evolving agriculture. The questionnaire is quite lengthy and detailed, although any given respondent should not have to complete all sections of the form. The COA asks extensive detail, and takes time and effort to read and report a farm operation. The use of one unified long form for all agricultural operations with all types of commodities leads to the perception of high response burden even if the real burden is less. Taken together these factors are suspected as contributing to declining unit response, underreporting at the commodity level, and item nonresponse. NASS is committed to reducing response burden. There are a number of ways to reduce response burden including: asking less questions, reducing detail collected in sections, asking sets of questions to only subsamples of respondents, reducing the sets of questions to answer, and reducing cognitive and perceived demands. To reduce respondent burden, both perceived and real, NASS would like to create shorter versions of the Census form, tailored to types of operations, for use during the 2017 Census, with selected groups of farm operations. There are several ways to reduce the number of pages on the questionnaire, including removing content, formatting differently, and including instructional material in different places on the questionnaire. One overall goal with any form changes, at a minimum, is to maintain response levels and maintain data quality. Any reduced questionnaire needs to be tested against a longer form for evaluating data reporting. In preparation for the 2017 Census, NASS worked with Washington State University to test the impact of shortening, redesigning elements, and reformatting forms.

In 2014, to design and test a newly formatted short form, NASS entered into a cooperative agreement with WSU to develop a condensed paper version of the 2017 Census of Agriculture. The short forms developed and tested were 16 and 12 pages, and varied in formatting. The 12-page questionnaire combined commodity code referencing into one comprehensive list which was displayed in one location on the questionnaire, and commodity item reporting was completed into one table. Thus this reduced the number of lines used by the respondent to report multiple commodities. The 16-page questionnaire had parsed commodity code item references for each type of commodity reported immediately preceding the questions for a particular commodity and had many more blank data lines used to report commodities. Both forms in the livestock section prelisted three most reported livestock categories. The 12-page questionnaire is the one that NASS is interested in using, if data for the two forms are comparable. The 12-page short form is

referred to as the SSF Blue 12-pg. The 16-page form is also a short form compared to COA long form, it is referred to as the SF Green 16-pg.

One mail out test was carried out for the purpose of comparing how a shorter paper form compares to similar slightly longer form and to gain information towards how best to collect agricultural operation information with short forms if used with selected portions of the farm population during the 2017 Census of Agriculture. Conducted by Washington State University (WSU), this test transpired in early 2015 to test the 12- page Reformatted Short Form against the 16-page similar form.

3. Sampling

Having previous data for farm operations from the NASS list frame is critical to evaluate both methods used to reduce questionnaire length. A key component of methods for questionnaire reduction and testing is the use of a pre-specified universe of farm operations that receive the shorter questionnaires. For the SFF Blue 12-page form, because we reduced the number of lines for commodity tables, we wanted to establish a threshold for the number of commodities an operation reported in past NASS data collections. To accomplish this, the list frame was targeted to select operations that received the forms in this test.

Content for the questionnaires was decided on in conjunction with the sampling universe planned for the short form. Using list frame data, operations with specific criteria and commodities were excluded from the sample for the test. The sampling universe was created by starting with the entire NASS list frame and then applying exclusion criteria. The very largest farms in each state, as identified using Total Value of Production (TVP) and Land in Farms (LIF), were eliminated using state level thresholds for these two variables.

In addition to removing the largest operations, operations classified as aquaculture, floriculture, Christmas tree, or organic farms were excluded, along with operations that grow certain specialty commodities. Finally, operations who grew more than 4 types of livestock, 4 types of poultry, 3 types of vegetables, 3 types of fruit/nuts/berries, or 6 types of field crops, or 10 total commodities were excluded. Most of these criteria were necessary because the reformatted short forms were made shorter by eliminating content related to specialty commodities and the number of commodities raised. Before sampling, the universe was sorted by state, county, farm size, and farm type. Systematic random sampling was used to select a sample of 6000 records which were divided randomly into two samples of 3000 cases each. After sampling was complete, regional offices were given the opportunity to remove cases from the sample based on prior arrangements with the operations. A total of 205 operations were removed, 99 from one sample and 106 from the other.

4. Questionnaires

The Census of Agriculture (COA) questionnaire used for the majority of the farm population (the “long” form) in 2017 will be approximately 24 pages and will be similar to long forms that have been used in past Censuses. The 2017 COA will incorporate some content changes, and possibly some design changes. The “short” questionnaire(s) are intended to be used for special populations with the aim of increasing response by providing them with a form that takes less time to fill out. Examples of this strategy were previously used in 2012, with several regional forms developed and used which collected

the same information, but had slightly different versions of matrix tables for collecting data on crops. In 2007, there were also regional long forms, along with a "short" form. The "long" and "short" versions collected the same content, but were formatted differently.

The last time NASS used a COA short form questionnaire was for the 2007 Census of Agriculture. No short form questionnaire was used for the 2012 COA, but there were updates made to the format and content of the 2007 questionnaire for the 2012 questionnaire. The 2012 COA form and the 2007 COA short form were used as starting points for the development of a short form questionnaire for the 2014 test. Washington State University (WSU) started with NASS's 2007 short form and enhanced some of the design features, creating a 16-page questionnaire, and for the purposes of comparison is called the SF Green Form 16-pg. The second short form developed was a 12-page reformatted short form referred to as the "Super Short Form" (SSF Blue 12-page) that combined the commodity codes into one list and reduced the number of lines in each question matrix table asking about commodities. The major differences between SF Green 16-page and SSF Blue 12 Page forms was the shortening of the questionnaire from 16 pages to 12 pages and this is one of the most significant changes made. Bringing together all the commodity codes, comprehensively, and centrally locating them in a "Commodity and Unit Code Table" on page 4 of the reformatted Short Form (SFF Blue 12-page) was another major design change. The number of lines in each table was reduced on the 12-page Reformatted Short Form (SFF Blue). In contrast, the 16-page form (SF Green) has shorter, very specific and separate commodity unit code listings located with each question in the matrix table. Other design changes included the SSF Blue 12-page form had accentuated vertical column grouping, visual highlighting, and instruction formatting. Additional graphical features (boldness, brightness, font size, capitalization, underlining, text color, less prominence of key codes were included in the SSF Blue 12-page form.

5. Methods

Ultimately, for this survey test, the shortened questionnaires were designed and developed to identify the most needed yet reduced set of information on agriculture required in the Census of Agriculture. For both questionnaire versions and this test, participants responded to questions corresponding to their use of agricultural land, their farming or ranching circumstances and operations. A principal method for the survey implementation was that respondents were contacted multiple times, in a single mode postal mail survey.

It should be noted that the implementation of this test as carried out by WSU, varies from NASS implemented surveys. The 12-page Reformatted Short Form (SFF Blue) was tested in a mail out compared to the 16-page Short Form (SF Green). In this survey test, WSU, under a collaborative cooperative agreement, was responsible for developing the questionnaires, as well as printing all materials and preparing them for mailing. NASS selected the sample respondents and produced numeric codes for tracking. SESRC staff were blind to the sample respondents. All questionnaire printing, mail assembly, processing, and data entry were handled by SESRC. All materials were bulk shipped to the NASS's NOD, and labeled and mailed by NASS's NOC. This was done to ensure that names and addresses of respondents remained confidential to NASS. Four sequential mailings using these procedures were sent to potential respondents--a presurvey letter, the first questionnaire packet, a reminder/thank you postcard, and a final second questionnaire packet to non-respondents. The presurvey letter, sent to all respondents, provided notification that a survey would be sent to them and that this was a collaboration of NASS and WSU to test the questionnaire for the Census of Agriculture. It also provided

information that the Census of Agriculture is required by law but that their participation in this trial run was voluntary and their responses are kept confidential by Federal law (Title 7, U. Code). WSU call center trained staff coded and flagged returned questionnaires. A web entry system with automated question branching was programmed and used to capture data entered by trained clerical staff. Respondents were informed that participation was voluntary and responses were ensured of privacy and confidentiality.

6. Analyses

For the overall reporting on the two versions of the reformatted short form, comparable response rates were achieved. The SF Green 16-page form stated with an initial sample of 2,901 and obtained 1,007 completes or a 35.6% response rate. For the SSF Blue 12-page form, the initial sample was 2,894 and received 991 completes and 5 partial completes with a response rate of 35.4%. The SF Green 16-page form had the more standardized appearance of routinely formatted NASS questionnaires. The unit level response was not significantly different between the two forms even though one form was 4 pages less in length. Table 1 shows the demographic breakdown of farm operations by operator age, race, and gender from the sampling frame and those who completed SSF Blue, 12-pg and SF Green 16-pg forms. The sampling frame consisted of 6,000 sample units, or operations. Of those, 37.25% did not have information regarding age of the operator, 9.42% did not have information about sex of the operator, and 9.67% did not have records about race/ethnicity. These cases and were not included in the analyses.

Farm operators in the sampling frame as well as the sample respondents for each of the SSF Blue 12-page and SF Green 16-page forms were very similar demographically, if not identical, in terms of their age distribution, race and gender. One concern for any form redesign in terms of data quality for agriculture is maintaining farm operation coverage in data reported and that data remains consistent across farm size as measured by Total Value of Production (Table 2) and the reporting of types of commodities. In this redesign, the shortening and reformatting of the forms did not distort or significantly impact the composition of operations reported for the survey, shown in Table 1 and Table 2. Operators aged 60+ comprised more than two-thirds of the sample, and the same percent of 60+ operators completed SSF (blue) and SF (green) forms. Those aged 40-59 comprised the next largest group of operators, accounting for 27.1% of farm operators. The remaining 2% of operators in the sampling frame were 18-39 years old. Majority (96.7%) of operators were white and male (86.8%). For Total Value of Production, the coverage across size categories is very similar with minor differences, i.e. the smallest size group is close to 8%, and the largest size group is 2% of responses for both forms, and this compares well to sample frame.

Table 3 shows the number of commodities reported and Table 4 shows the major commodities groups reported by agricultural operations in the sample frame and for both of the shortened forms. These results and statistical tests confirmed the two forms performed comparably for the percentage of operators reporting the number of commodities produced (Table 3) and percentages reporting in the major groups of commodities (Table 4) with no disparate differences between them in coverage of commodities.

6.1 Analyzing Data Quality at the Item Level

Data quality can be explored in a number of ways and the ways to evaluate this quality are associated with the types of questions and the types of response options. The Census of Agriculture is a complex detailed questionnaire and the reporting burden and difficulty of answering increases with the combinations of operation size, complexity of commodities, and the extensiveness of agricultural operations. The reporting burden and difficulty is further compounded with the intersection with the form of the question used to obtain answers, item definitions and instructions, and questionnaire aids such as instructions and commodity lists. There are a number of types of questions included in the COA short forms—screening questions, categorical questions, numeric answer questions, and matrix tables of questions that combine open-ended write-in answers with numeric and categorical responses. In some instances, questions require mathematical calculations and equivalency comparisons for respondents to check their reporting is accurate. The types of questions include screening questions with categorical answers at the beginning of Sections, categorical questions throughout the form, open-ended write-in questions, and questions with numeric response or number codes. The purpose of screening questions is to prevent item response and for respondents to help branch or skip past questions that don't apply. Numeric questions have been developed that allow for a numeric answer or the checking of a "none" box. An overarching concern in agriculture reporting is to obtain detection of when specific types of agriculture production is occurring and when it is not, as agricultural land can go in and out of production over time. Each of these types of questions are associated with specific types of reporting errors. Data quality was assessed in this study by evaluating: the number of questions answered, the lines of data provided, the number of questions where numeric information was provided, numeric comparisons of averages for commodities, the use of check boxes, zero answers, missing items, and the performance of screening questions to determine if there were differences between the two forms.

One visual design difference between the two forms occurred in the formatting of the "None" Check boxes intended for use by respondents to report where there is zero or a none response is shown in Figure 1. The purpose of this type of response option is to reduce "missing" answers to questions. The visual design of this feature for the SF Green 16-Page form incorporates a highlighted white space around the outside of the lined check box to facilitate the optical scanning of responses by NASS. The design on the SFF Blue 12-Page Form does not have this feature. To determine if this had any effect on the quality and data collected, the percent of respondents who checked the "None" box was compared across questions that had this format for responses across the two forms and is shown in **Table 5**. In addition, **Table 6** also displays results for a selected subset of questions on the provision of positive or "greater than zero" response, write in of "zero or 0", and missing blank answers. With provision of a "None" check box on the form, quality or accurate answers in numeric answer spaces, would be the checking of the None box, or answering with a number greater than zero. A written zero provides an answer that needs data cleaning and handled more at the operational level. The provision of a none box and use by respondents allows for optical scanning and lessens data cleaning. The most inaccurate answers are missing responses---left as blank. As shown, in table comparisons (Tables 5 and 6) at the item level, the visual design led to significant differences in respondents' use of the "None" check boxes with the SF Green 16-Page Form having a higher percentage of use than in the SFF Blue Form. Looking further at Table 6, the SF Blue 12-page form "check box" design (without the highlight), led to a higher percentage of respondents giving a zero response across questions. The SF Green 16-page form had more missing responses. For the reporting of agricultural information that contributes to numeric estimates, that is

greater than zero responses, both forms had comparable responses on the percentages of respondents reporting >zero, positive answers, on items. Significant differences were associated with item missing and write-in of zero.

6.2 Interaction of Visual Design, Data, and Errors in Reporting

For commodity reporting matrix style tables were used and between the two forms there were slight variations in design of commodity matrix table sections. For example, Figure 2 shows the livestock reporting pages in the two forms. The SF Green 16-page form visually had more lines for reporting commodities in each table and also had a reduced list of specific commodities and their reference codes located within the table. By comparison, the SSF Blue 12-page form used a smaller number of reporting lines in each matrix and the commodity code reference list was provided in one centralized comprehensive table located at the start of the commodity reporting sections. This required respondents to refer back to the list as needed. Another difference was the presentation of the commodity reference code reporting column between the forms. The SSF Blue 12-page form used 4 segments in the cell in each row of a commodity to report a 4-digit commodity code. We discuss the influence of these designs and considerations.

Further evaluation of numeric data yield at the commodity level found similar reporting of the number of crops and commodities and lines of data provided between both forms with no significant differences, statistically. We confine our discussion to field crops and livestock. First, looking at field crop reporting, the longer SF Green 16-page form had slightly more operations reporting field crops, 340 operations compared to 283 for the SSF Blue 12-page form. Field crop reporting also showed slightly more with 33.8% versus 28.6% of respondents reporting a field crop, on the SF Green 16-page form compared to the SSF Blue 12-page form. The SSF Blue 12-page form had slightly less lines of data compared to the SF Green 16-page form at 386 lines of data for 283 operations compared to 434 lines of data for 340 operations. The average number of field crops reported on the SF Green 16-page form was 1.28 crops, and 1.36 crops on the SSF Blue 12-page form.

Of the total number of operations more than a third, 34% or 677 operations reported at least one line of livestock data. We found no statistical differences between the two forms for livestock reporting and the 3 prelisted items of the most often reported livestock—beef cattle, milk cows, and other cattle. For the SSF Blue 12-page form there were 521 lines of data for 336 operations. For the SF Green 16-page form there were 562 lines of data for 341 operations. For the leading livestock category reported, beef cows, 51.4% respondents for the SSF Blue 12-page form reported this category compared to 48.4% for the SF Green 16-page form. Figure 3 displays a graph of the amount of livestock reported for the two forms, with no statistical differences. However, the longer Green 16-page form with more lines in commodities tables and the location of the specific livestock commodity code list above the table for reporting captured slightly more reporting of other types of livestock.

For each type of agriculture commodity produced on an operation, information about the total level of production is entered into matrix grid style tables; that are physically separated for each type of commodity group in a section of the form. Matrix tables (Figure 2) are burdensome for respondents as they have to attend to multiple types of alpha and numeric information to be reported and multiple levels of cognitive processing to retrieve and categorize specific answers. Levels of information are to be reported sequentially in a row for each commodity. After the crop or commodity name, the table has a column cell

for reporting a commodity code. This is specific piece of information is unfamiliar to the respondent and requires finding the reference in the commodity code table provided in the form and then transferring this to the cell. Entering information into the commodity tables is associated with various types of errors. Figure 4 and Figure 5 summarize the types and extent of errors for field crops and livestock. Ten types of errors were recorded for field crop entry in the field crop reporting table. The most common types of error across both forms for field crops were: missing the crop code, incorrect crop code, and missing crop name. Another problem was respondents placing information in the incorrect column or merged with another answer and associated with field crops was the improper reporting of the unit of weight placed in the total quantity harvested column. There were 121 errors in 820 lines of data. In the livestock matrix table, less errors were found overall and this was associated with the provision of preprinted livestock names and commodity codes. The most prevalent errors for both forms were: missing the livestock code, missing the livestock category name, and entered in an incorrect row of the table. The SF Green form had slightly more errors in reporting in the livestock table. It is assessed this was associated with the visual layout differences between the two forms.

7. Conclusions

In this paper we have examined several comparisons of resulting outcomes of unit level response and item level response from testing associated with the use of a shorter reformatted form for Census of Agriculture. In terms of hypotheses tested, for the first hypothesis, we reject the null hypotheses that a shorter form will increase rate of response. A shorter form by four pages did not result in increased unit response. We found no measureable differences between the two short forms, with neither garnering a significantly higher level of response. A future test is warranted to compare the selected short form to the full 24 page COA for specially selected types of operations for unit level response. The current testing provides a benchmark. A shortened form potentially has an advantage for response compared to the full COA long form.

In terms of overall data quality assessment, one important confirmation from the testing is that the shorter SSF Blue 12-page form yielded information on agricultural production comparable to the 16 page longer form. The shortening of the form by four pages had no measure impact on data quality in terms of respondents reporting levels of actual production by entering numeric amounts (greater than zero) for items produced (Moore and Gertseva, 2016). The extent of reporting of numeric responses was comparable between the forms and we few measureable differences for the more reported items (Moore and Gertseva, 2016). Further analyses were completed at the commodity item level to asses if agricultural statistical measures were distorted. This is beyond the scope of the current paper.

The most promising redesign feature tested in this study was placement of the commodity unit code tables. An important aspect of commodity level data quality is the use of item unit codes by respondents as this reduces error in recording the name or specific type of commodity entered in a commodity table. Like others, Snikers et. al. (2013) and Perry (2007) we find question design to be key to collecting information and eliciting respondents to self-select accurate establishment codes. The testing confirmed the null hypothesis for the second hypothesis tested. Placement of the specific commodity unit code list at the

question reduced item nonresponse for unit codes in the field crop table where respondents were intended to hand enter the unit code. The SSF Blue 12-page form included the centralized comprehensive commodity unit code table and this resulted in a significantly higher rate of missing commodity unit code responses for field crops as compared to the SF Green 16-page form which showed a pared down unit code reference table that included only those items relevant to the specific question. For the livestock section, in both forms, the use of three preprinted main livestock commodities reduced the level of error in recording entries of type of livestock and livestock unit codes. However, there was more error in entry of other livestock associated with SSF Blue 12-page form, where respondents had to leave the page and flip back in the form to the centralized commodity unit code table. Going forward, we recommend placement of specific commodity unit codes at the question as a prominent design element for the COA short form.

Of the aspects of data quality examined for data collection for the two short form designs, one of the most notable inconsistencies in data reporting were associated with an unintentional test of a graphical element, the display of the "None" check box. The "None" check box feature is very important for data quality as on any given item a large portion of respondents may not have this characteristic or produce a commodity. For instance, almost 69% of respondent indicated no cropland had failed on the SF Green 16-page form. The results show that respondents are impacted by graphical elements, and specifically shows how respondents interacted with the form to specify no production---the use of "None" check boxes, writing in zeros in numeric fields, or leaving items as missing or blank. The SSF Blue 12-page form and its lack of highlights around the check box graphic led to a lower use of the "None" check box by respondents compared to the SF Green 16-page form. This design also led to a higher rate of respondents writing in zeros for questions with numeric fields. This type of data quality error has consequences for costs associated with data processing and cleaning. This type of change to the graphical display of "None" check boxes, if implemented, could potentially lead to increased costs as more cases would not be handled automatically with optical scanning, we highly recommend that NASS stay with its traditional design of highlighting the check box.

Extending research for evaluating the impacts of shortening questionnaires on data quality should be extended to other types of enterprise surveys beyond agriculture and this may be an important area of future research. The idea of parsing classification tables throughout a questionnaire as a way to reduce respondent burden and also as a way to reduce respondent selected data coding error for enterprise surveys should be explored in other contexts.

Figure 1. Visual Design of "None" Check Boxes in Blue 12-page to SF Green 16-page form.

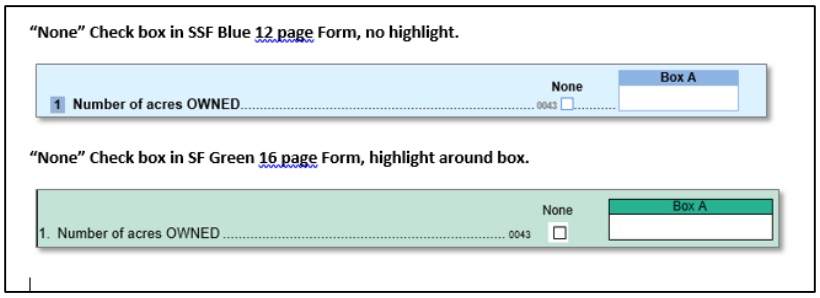


Table 1. Demographics Analysis for Sampling Frame and for Respondents by Form.

	Sampling Frame		SSF (Blue 12-Pg)		SF (Green 16-Pg)	
	N	%	N	%	N	%
Age of operator						
18-39	70	1.9%	5	0.7%	3	0.43%
40-59	1,021	27.1%	159	22.5%	135	19.2%
60+	2,674	71%	543	76.8%	564	80.3%
Total	3,765	100%	707	100%	702	100%
Race						
White	5,248	96.7%	890	96.4%	906	98%
Non-White	172	3.2%	33	3.6%	20	2%
Total	5,420	100%	923	100%	926	100%
Gender						
Female	716	13.2%	117	12.5%	119	13%
Male	4,719	86.8%	805	87.5%	808	87.2%
Total	5,435	100%	922	100%	927	100%

Table 2. Total Value of Production for Sample and for Respondents By Form.

Total Value of Production	Sampling Frame		SSF (Blue 12 Pg)		SF (Green 16 Pg)	
	N	**%**	**N**	**%**	**N**	**%**
$1-$999	494	8.2%	76	7.7%	81	8.1%
$1,000-$2,499	519	8.7%	87	8.8%	95	9.5%
$2,500-$4,999	737	12.3%	115	11.6%	112	11.2%
$5,000-$9,999	922	15.4%	151	15.2%	174	17.3%
$10,000-$24,999	1,349	22.5%	227	22.9%	232	23.1%
$25,000-$49,999	869	14.5%	165	16.7%	135	13.5%
$50,000-$99,999	561	9.4%	95	9.6%	99	9.9%
$100,000-$249,999	362	6.0%	56	5.6%	53	5.3%
$250,000+	181	3.0%	18	1.8%	22	2.2%
Total	5,994	100%	990	100%	1,003	100%

Table 3. Commodities Reported for the Sample and for Respondents by Form.

Number Commodities	Sampling Frame		SSF (Blue 12 Pg)		SF (Green 16 Pg)	
	N	%	N	%	N	%
1	2,150	38.5%	350	37%	370	39.3%
2	1,871	33.5%	318	33.7%	309	33%
3	1,155	20.7%	200	21%	193	20.5%
4	285	5.1%	63	6.7%	44	4.7%
5 or more	126	2.3%	15	1.6%	26	2.8%
Total	5,587	100%	946	100%	942	100%

Table 4. Type of Operation for the Sampling Frame and for Respondents by Form.

Farm Type	Sampling Frame		SSF (Blue 12 Pg)		SF (Green 16 Pg)	
	N	%	N	%	N	%
Grains	758	12.7%	119	12%	125	12.5%
Vegetables, Melons, Potatoes, Fruit, nut, berries	310	5.2%	46	4.7%	37	3.7%
Other crops	1905	31.8%	338	38%	353	35%
Livestock	2265	37.8%	388	39%	370	37%
Cropland only	756	12.6%	99	10%	118	11.8%
Total	5,994	100%	990	100%	1,003	100%

Figure 2. Comparison Livestock Matrix Table Pages,
SSF Blue 12-page Form Compared to SF Green 16-page Form.

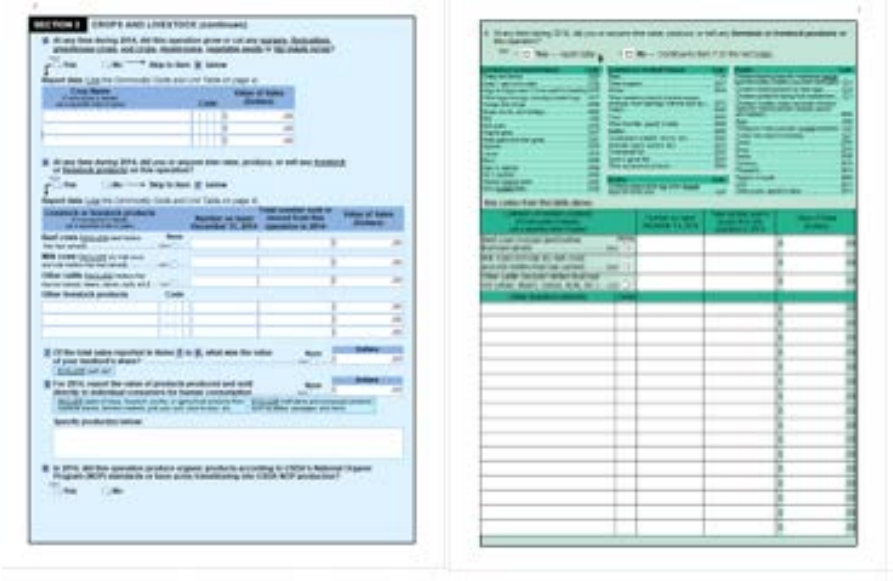


Figure 3. Percentage Reporting Livestock Lines of Data and Categories.

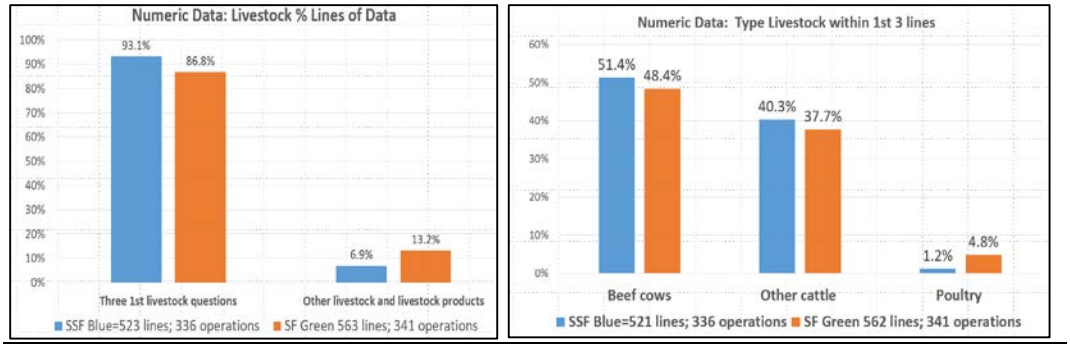


Table 5. Selected Questions, Respondents Checking the "None" Box, by Form.

Category	SSF Blue 12-Pgn-990	SF Green 16-Pgn=1001	χ^2	P-value
Acres Owned	3.6%	4.0%	0.17	0.67
Acres rented from	51.8%	64.8%	**34.7**	**<0.001**
Acres rented to	49.5%	60.0%	**22.5**	**<0.001**
Cropland harvested	36.4%	40.4%	3.36	0.067
Cropland failed	55.6%	68.2%	**33.92**	**<0.001**
Cropland summer	56.4%	68.5%	**31.43**	**<0.001**
Cropland not grazed or harvested	47.1%	56.4%	**17.51**	**<0.001**
Permanent pasture	27.5%	36.3%	**17.7**	**<0.001**

Table 6. Responses Selected Questions with Numeric and Missing Answers, By Form.

	SSF (Blue 12-pg) (N=990)						SF (Green 16-pg) (N=1001)							
	Zero		>Zero		Missing		Zero		>Zero		Missing			
Question	N	%	N	%	N	%	N	%	N	%	N	%	χ^2	P-value
Section 1: Acreage 2014														
Acres owned	38	3.8%	938	94.8%	44	4.4%	8	0.8%	949	94.8%	44	4.4%	0.003	0.99
Acres rented from	363	36.7%	139	14.0%	488	49.3%	285	28.5%	152	15.2%	564	56.4%	**15.4**	**<0.001**
Acres rented to	304	30.7%	220	22.2%	466	47.1%	221	22.1%	244	24.4%	536	53.6%	**19.2**	**<0.001**
How many acres	171	17.3%	180	18.2%	639	64.6%	122	12.2%	201	20.1%	678	67.7%	**19.2**	**<0.001**
Section 2: Land														
Cropland harvested	60	6.1%	423	42.7%	507	51.2%	34	3.4%	391	39.1%	576	57.5%	**12.8**	**0.002**
Cropland failed	131	13.2%	38	3.8%	821	82.9%	80	8.0%	23	2.3%	898	89.7%	**19.4**	**<0.001**
Cropland summer	127	12.8%	18	1.8%	845	85.4%	68	6.8%	22	2.2%	911	91.0%	**20.7**	**<0.001**
Crop land not harvested	113	11.4%	138	13.9%	739	74.7%	54	5.4%	158	15.8%	795	79.4%	**23.5**	**<0.001**
pasture	71	7.1%	434	43.8%	485	49.0%	36	3.6%	426	42.6%	539	53.9%	**14.3**	**0.0008**
Woodland pasture	112	11.3%	182	18.4%	696	70.3%	59	5.9%	152	15.2%	790	78.9%	**25**	**<0.001**
Woodland not pastured	91	9.2%	306	30.9%	593	59.9%	45	4.5%	310	31.0%	646	64.5%	**17.8**	**0.0001**
All Other land	77	7.8%	410	41.4%	502	50.7%	24	2.4%	465	46.5%	512	51.2%	**32.3**	**<0.001**
Total acres	9	0.9%	787	79.5%	194	19.6%	12	1.2%	826	82.5%	163	16.3%	4.003	0.135

Figure 4. Qualitative Analysis of the Types of Errors for Field Crops by Form.

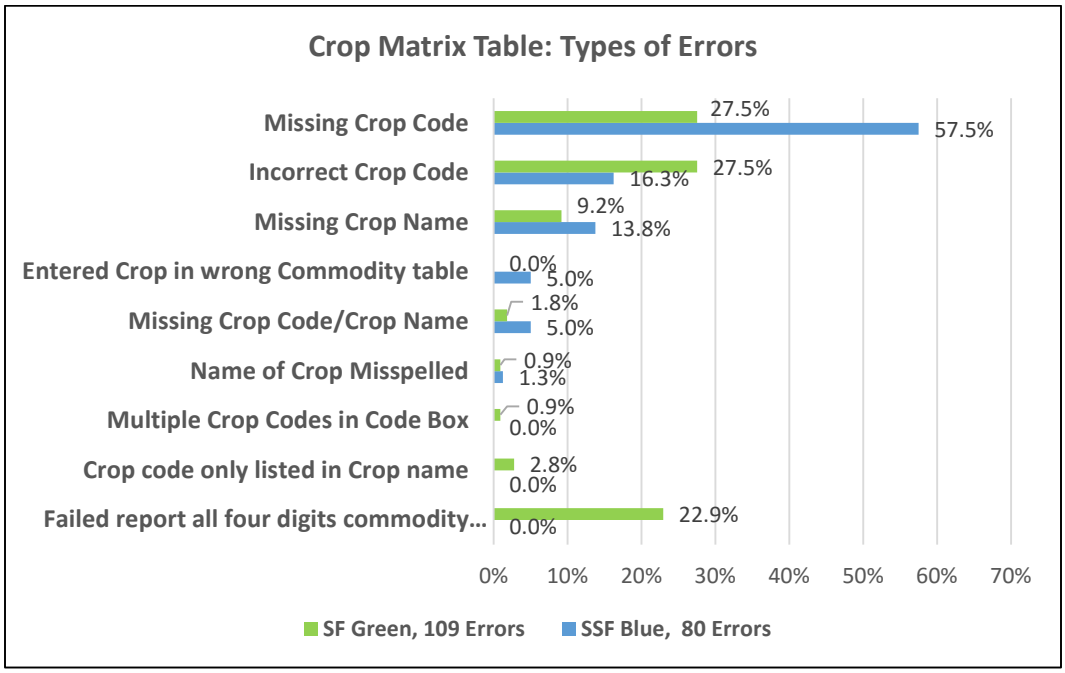
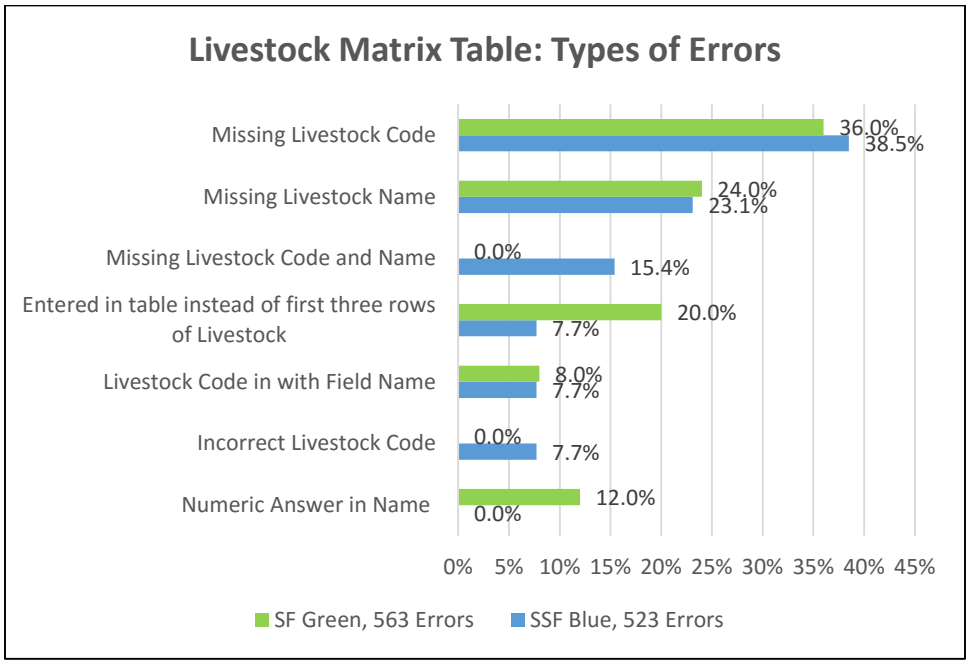


Figure 5. Qualitative Analysis for Error Types for Livestock by Form.



References

Axhausen, K.W. and Weis, D. (2010). Predicting Response Rate: A Natural Experiment. Survey Practice. (3) 2. www.surveypractice.org. ISSN:2168-0094.

Bruvold, N.T. and J.M. Comer. (1988). A model for estimating the response rate to a mailed survey. *Journal of Business Research 16*(2): 101–116.

C-FARE. (2007). Improving Information About Americas Farms & Ranches: A Review of the Census of Agriculture. Report of the Council on Food, Agricultural and Resource Economics (C-FARE) Washington, D.C.

de Leeuw, E. and Collins. (1997). Data Collection Methods and Survey Quality: An Overview. Chapter 8. Survey Measurement and Process Quality, Edited by Lyberg, Biemer, Collins de Leeuw, Dippo, Schwarz, Trewin. John Wiley & Sons, Inc.

Dillman, D.A., Smyth, J.D., and Christian, L.M. (2014). Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method. Wiley & Sons, Hoboken, New Jersey.

Fox, R.J., M.R. Crask and J. Kim. (1988). Mail survey response rate: a meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly 52*(4): 467–491.

Groves, R.M., E. Singer and A. Corning. (2000). Leverage-saliency theory of survey participation: description and illustration. *Public Opinion Quarterly 64*(3): 299–308.

Groves, R.M., S. Presser and S. Dipko. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly 68*(1): 2–31.

Heberlein, T.A. and R. Baumgartner. (1978). Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature. *American Sociological Review 43*(4): 447–462.

Moore, Danna L. and Gertseva, A. (2016). Developing and Testing the 2017 Census of Agriculture Form: Testing the 2017 Census of Agriculture Condensed Form Against the 2007 Census of Agriculture Short Form. Technical Report 16-56 of the Social and Economic Sciences Research Center, Washington State University.

National Research Council. (2008). Understanding American Agriculture: Challenges for the Resource Management Survey. Panel to Review USDA's Agricultural Resource Management Survey. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, D.C: the National Academies Press.

Perry, J. (2007). Implementing coding tools for a new classification. Proceedings International Conference of Establishment Surveys (ICES 3). Montreal, CA. June 28-21, American Statistical Association, Alexandria, VA, pp. 33-41.

Snikers, G., Harldsen, G., Jones, J., and Willimack, D.K. (2013). Designing and Conducting Business Surveys. Hoboken, New Jersey: Wiley.