

**U.S. DEPARTMENT OF AGRICULTURE
FOOD AND NUTRITION SERVICE**

***Enhancing Completion Rates for SNAP
(Supplemental Nutrition Assistance Program) Quality Control Reviews***

*Request for Clearance
Supporting Statement and
Data Collection Instruments*

*Attachment C.1:
Review by the USDA National Agricultural Statistics Service (NASS)*

Project Officer: Robert Dalrymple

September 26, 2013

ATTACHMENT C.1: REVIEW BY THE USDA NATIONAL AGRICULTURAL STATISTICS SERVICE (NASS)

REVIEW OF THE ENHANCING COMPLETION RATES FOR SNAP (SUPPLEMENTAL NUTRITIONS ASSISTANCE PROGRAM) QUALITY CONTROL REVIEWS

Michael Jacobsen
June 14, 2013

Part A – Purpose and Justification

The justification and the purpose for this study are inadequate. You state that the purpose of this study is to do two things: 1) describe a typical incomplete QC review and 2) figure out how to improve QC review rates. You then justify this by stating that QC completion rates have undergone a dip in the past 28 years and that state QC completion rates greatly vary. This does not seem to be an adequate purpose for conducting a national survey. The first “purpose” could be easily handled by an analyst in FNS. The second “purpose” could be handled by a discussion between state SNAP Directors at a national conference.

The justification in part A.1 lacks substance. First, you mention that a random sample is drawn for QC review. However, no details of the sample are given. What is the size of the sample? What is the sample design? How much money and how many resources are needed to conduct reviews on all sampled cases? (i.e., what is the “cost” of the sample?) Is auxiliary information used to design the sample? What is the purpose of this sample? When I ask the last question, I want to know what this sample of households intends to accomplish in terms of estimating validity and payment error rates. For example, are you looking for the number of households improperly receiving SNAP benefits to be 10% with a 95% confidence interval of 8% and 12%? (If you did not already know, a 95% confidence interval is the range that an estimate will be between in 95% of the time if the same sample is drawn multiple times.) By describing the random sample, you can discuss the actual issue addressed in your project.

Second, you offer this statement: “Statistical rigor requires completing as many reviews as possible.” This is a vague sentence. What do you mean by “statistical rigor?” Are you talking about increasing the number of QC review cases? If so, obtaining more observations past what the sample requires has a diminishing effect on improving the estimate. Are you talking about decreasing the incomplete and refusal cases? This would make sense given the purpose of this project.

Third, you mention that SNAP QC case completion rates decreased between 1985 and 2006 and then rose somewhat after 2006. What caused the change in rates? How big were the changes? Why were they large enough to justify a nationwide review of QC case handling? Has the rate of increase since 2006 been increasing, decreasing or staying the same? Furthermore, you mention that the QC completion rates vary among the states. Do all of the states show the same trend or is the national trend being driven by a few states? Why is this change important?

Fourth, another issue with the time frames is that there are three different time periods mentioned in part A: 1985 to now (the overall decrease in completion rates), 1985 to 2006 (the unchecked decrease in completion rates) and 2006 to now (the increase in completion rates). Which time frame are you studying? How will you distinguish between answers that are for the 1985-2006 time frame versus those for the 2006-now time frame versus answers for the overall time frame? Furthermore, is the distribution of incomplete and refusal cases the same in each time period? Have the number and type of QC cases remained constant in all three time periods or do they differ? Does the number of QC cases differ with time periods?

After reviewing this docket multiple times, it seems to me that you want to measure the effect of incomplete and refusal QC cases on the ability of the randomly drawn sample to estimate the validity of SNAP cases across the population of households receiving SNAP benefits. If this is true, the justification would then be something like this: QC reviews cost money that should be going to giving SNAP benefits to deserving households. However, QC reviews are needed to see if the households receiving the benefits truly deserved them. There are not enough resources to review every case; therefore, a sample is randomly drawn from all cases of QC review. Since 1985, the completion rate of sampled QC reviews has decreased nationwide, reaching a low in 2006.

While the completion rates have increased somewhat since 2006, you do not know how well these additional completed cases represent the SNAP recipient population validity estimate. Furthermore, you do not know the distribution of the incomplete and refusal cases. Therefore, you do not know whether your estimate of validity is affected by potential biases due to the incompletes and refusals. If the validity is affected, then your estimate of the payment error rate is unreliable and the SNAP money may not be appropriately distributed despite the recent increase in completion rates.

To address this issue, you could increase your sample size to match the number of potential incomplete and refusal cases or you could examine the characteristics of the incomplete and refusal cases. Increasing the sample size is less desirable because it shifts money and resources away from SNAP benefit administration to SNAP QC review and it does not explain the reasons behind the increase in incomplete and refusal cases. Examining the incomplete and refusal population is optimal because you can adjust or redesign the sample to address the incompletes and refusals and you will better estimate the SNAP validity due to the increased knowledge about SNAP QC incomplete and refusal cases.

The research objectives in part A.2 could then be condensed into three specific objectives:

1. Measure the effect and the potential bias of the incomplete and refusal cases on the validity estimate (original objectives 3, 6, and 7).
2. Measure the effect of and potential bias due to the processes of conducting QC reviews at State and Federal regional levels on the validity estimate (original objectives 1, 2, and 5).
3. Identify the challenges and best practices to reduce the impact of the bias on the validity estimate (original objective 4).

You can meet these objectives using the four tools you have outlined in part A.2. Specifying a time frame will also be necessary.

Part A – Other Questions

In part A.1, you mention that approval for the QC re-reviews on SNAP participants was previously granted by OMB. In a footnote, you say that the approval included a review schedule with one OMB approval number and a worksheet with another approval number. First, are the worksheet and review schedule the only items you need for the re-reviews? The footnote implied that you would be using other items. Second, will the SNAP participants be asked for any new information in the re-reviews that were not asked for in the previously approved questionnaires? If no new information will be collected, then you are correct in saying that the previous OMB approval applies to the re-reviews. However, if the SNAP participants will be asked for information that was not included in the previously approved projects, then the re-review materials need to be included in this docket and their burden calculated.

In part A.8, the public comment notice was published in which volume of the Federal Register? There are spaces in the paragraph for the volume, number, and page numbers, as well as a reminder to add a statement on receiving no public comments.

In part A.12, what is the burden on the State SNAP offices to collect the administrative data? I assume that the admin data will be coming from the State offices; therefore, there should be a burden estimate for the State offices.

In table A.12.1, the fourth column from the right (the “Total Annual Response” column) has some strange subtotals. For example, the State QC Directors Subtotal is 80. What does this 80 mean? It cannot mean the number of non respondents after data collection has ended; the column contains the number of non respondents after that particular data collection effort. Therefore, the subtotal should be 10. It also cannot mean the number of non responses by each non respondent at each stage of data collection. Each non respondent only had five chances to not respond. Furthermore, this 80 is greater than the starting sample of 53. So, what does this subtotal represent? I have the same questions for the State QC Supervisors and Reviewers subtotals.

The “Number of Respondents” column in table A.12.2 is improperly labeled. The column title should be “Sample Size” or something similar. The values in the column are the sample sizes for each of the subgroups, not the number of respondents for each subgroup. (Incidentally, the “Type of Respondents” column is also improperly labeled. It should be “Subgroup” or something similar.)

Part B

In part B.1.2, you mention that the survey frame will include reviewers, supervisors and directors that have recently left their positions. Why would you want these people in your frame? I ask this for two reasons. First, the term “recently” is too vague to mean anything. Second, the longer a respondent is removed in time from a situation, the more unreliable their responses become and the more survey error is introduced into your findings. Third, in Attachment E, you discovered that active reviewers who responded to the pretest could not remember numbers and estimates for cases three months prior to the pretest. Even though you rewrote the questions to compensate for this, these are reviewers who are still at their jobs and they cannot remember the details. How do you expect someone who has been away from QC review to remember the answers to you reworded questions? I would not recommend including these people on your frame.

In part B.1.3, how did you come up with the response rate estimates of 81 percent, 82 percent and 83 percent for the directors, supervisors and reviewers, respectively? There should be some indication of the validity of these estimates mentioned here (e.g., a previous study or the results of a test study). Also, do you expect you non respondents to be randomly selected? In other words, could the non responding reviewers/directors/supervisors bias your survey findings away from the true results? Any kind of prior knowledge about the observations in your survey frame would greatly help to address this issue.

In part B.2, you mention research showing that the increase in response rates decreases beyond seven calls but you do not cite the source. Where did these findings come from? I think that you are referring to the 2001 Groves, Wissoker, Greene, McNeeley and Montemarano study mentioned in *Survey Methodology*, 2nd Ed. This study showed that the percentage of households first contacted for a survey decreases as the number of calls to the non contacted households increases. However, you may have a different study in mind. If so, please cite it.

How are you going to summarize and analyze the data from the semi-structured interviews? You mention in part A that the answers will be coded and analyzed in NVivo v10 software but no more information was offered beyond this point. Furthermore, no information on the NVivo analysis is mentioned in part B. More information on the NVivo coding procedures and analysis would be helpful in describing the outcomes of the semi-structured interviews. Furthermore, you are not going to try to combine the data from the semi-structured interviews with the data from the web and phone surveys, are you? It did not seem that you would combine the estimates but I am still unsure after reading the docket.

Questionnaires – Overall Issues

There are several recurring issues with the questionnaires that, if left unaddressed, will increase the amount of nonsampling error in the final estimates. The first issue is that time frames differ between questions. One question may ask for an answer for the previous month while another question requires an answer for the respondent's entire time working on QC cases. The problem is that the answer to one of these questions may have nothing to do with the answer to another question. Consequently, you may try to identify connections between responses that are not real.

The second issue is the use of vague and ambiguous words and phrases in the questions and supplied answers. Respondents may use different interpretations when reading these words; therefore, their responses may vary too much for your estimates to be of much use. If they occur in the semi-structured interviews, the interviewer will have to probe more and will increase the chance of introducing interviewer error into the response.

The third issue is asking questions that test the respondent's knowledge about a particular subject and introduces a social desirability bias. The problem is that this may not be the data about the respondent that you want. For example, the respondent may have learned the most effective procedures for QC reviews from some authority but uses other procedures that are the most effective for his or her reviews. However, when you ask about the most effective procedures for QC reviews, the reviewer can legitimately tell you the procedures from the authority and not the ones he or she actually uses. Therefore, your data will not be accurate.

The fourth issue is the lack of nondirective probes for the phone and face-to-face interviewers. While you can write questions that are logical and precise to you, the respondent will most likely not understand the questions or the required responses or will say something that does not answer the question. Consequently, the interviewer will have to probe the respondent to get the correct answer for that question. The problem arises when the interviewer does not understand the question or probes in such a way that leads the respondent to an answer. In both cases, interviewer error is introduced and the data may not be correct.

The fifth issue is asking multiple questions in one question. This issue tends to be limited to the semi-structured interviews but it runs the risk of overtaxing the respondent's cognitive processes in answering questions. Therefore, the respondent may not give an accurate answer because he or she is processing all of the questions at once. Furthermore, interviewer time and error will increase because the interviewer may need to probe more to have the respondent answer both questions.

State QC Reviewer Semi-Structured Interview

- Question 3 asks about two different time frames: five years versus month-to-month. What is the final estimate you are trying to find? Are you looking for a net change in cases over the past five years or are you looking for monthly fluctuations? What is the difference between “stays the same” and “change month to month” if the net change is zero with monthly variance?
- Question 4 contains multiple problems. First it asks multiple questions at once. The individual questions should be separated for ease of answering. Second, the second question is a point-of-view question that may not reflect the actual distribution of the QC cases by the supervisor. Is an opinion what you are seeking with this question? Third, “fairly” and “frequently” are ambiguous and can be too broadly interpreted.
- Question 5 also asks multiple questions that should be split into two individual questions. Furthermore, to which question is the question “How often?” referring? “Often” means what?
- What does “homeless” mean in Question 10? Different people have different meanings for “homeless.”
- Question 12 asks multiple questions that should be separated.
- Question 13 asks multiple questions that should be separated. Furthermore, “times of day” and “days of the week” could mean a period such as “morning” or “midweek,” a count of time within one of the previously mentioned periods, or an actual range of time. What type of answer do you want?
- Question 14.B asks for the interim deadlines given throughout the QC process. However, this question seems very broad and case-dependent. Consequently, the answers you receive may reflect a range of cases and may not offer any general information.
- Is Question 16 the same as Question 14? It would seem that way, especially if the respondent answers in days on Question 16. What is the difference between the two questions? Also, what “average” are you referring to? Cases? Hours? Days? How will you differentiate between the answers?
- Question 17 asks multiple questions that should be separated. Also, the first question asks for one of two different types of answers. The respondent must answer “yes” or “it varies” which does not mean “no.” Furthermore, what would respondents interpret as following the steps outlined in the handbook? This could be interpreted as 1) none of the steps are followed; 2) all of the steps are followed; 3) only some of the steps are followed; and 4) all of the steps plus additional steps are followed. 1 is a “yes” answer and 2 is an “it varies” answer. However, 3 and 4 would have both responses as valid answers, which would muddy the inference. What type of answers do you want for the first question?
- Question 18 asks multiple questions that should be separated. The first question is very confusing and needs to be reworded to be similar to the second question.
- What do you mean by “general steps” in Question 20? Are you referring to the steps taken in an average case? Are you referring to the respondent’s preferred steps? Are you referring to the absolute necessary steps for all cases?
- What is the time frame for Question 20.A? Is it the previous month as asked about in Question 1? Is it the five years mentioned in Question 3? Or is it neither of those? In any case, there is no confidence that the answer to this question has any relation to the answers to Questions 1 and 3. In other words, you cannot make any connections between the information provided by Questions 1 and 3 and the information from Question 20.A because of the different time frames.

- What do you mean by “attempt” in Question 20.B? Will the respondent have the same definition as you?
- What does “average” refer to in Question 20.C? Average day? Average case? There could be too many answers for this question to be of any use.
- Why is another time frame introduced in Question 20.C.1? Again, you cannot relate these answers to the answers to previous questions.
- The wording of Question 20.D could introduce a social desirability bias into the data. Respondents could answer “the most effective strategy” as the strategy promoted by the State and Federal offices as the most effective. However, this does not mean that it is the respondent’s most effective strategy. If you want the respondent’s most effective strategy, I would recommend changing “the” to “your.”
- Question 21 and its parts have the same issues of different time frames (Questions 21.A and 21.B.1), social desirability bias (Question 21.D), and ambiguous phrases (Questions 21 and 21.B) as Question 20. Any corrections used in Question 20 should be applied to Question 21.
- In Question 21.E, how will you probe the respondent for further information? What type of nondirective probes will the interviewer use to both obtain the desired response and suppress the amount of interviewer error introduced?
- Question 23 asks multiple questions that should be separated. Also, social desirability bias is introduced in Question 23. Respondents should be asked for their strategies, not the strategies.
- What time frame are you referring to in Question 27.A?
- Question 28 is not a survey question. It is a test question that measures the respondent’s knowledge on the proper way of conducting a field interview. Change “taken” to “you take” to have the respondent provide their actual answer.
- Question 28.B is another test question. Ask the respondent if he or she takes the same steps for every case.
- Question 28.C is another test question. Ask the respondent if he or she has changed these steps. Furthermore, the time frame defined in Question 28.C is very broad and a QC reviewer could have seen multiple changes in the steps taken. Do you want all of the step changes, even though the respondent may not use those changes anymore? Or are you looking to see the changes in steps taken, say, between 2006 and now? What answers are you expecting?
- Question 29 has four issues. First, it is a test question and not a survey question. What if the State office recommends conducting interviews in the morning but the respondent has the best results conducting the reviews at night? The respondent could respond that the mornings are better but that is not the desired answer. Reword the question to ask respondents for their optimal times. Second, it asks multiple questions that should be separated. Third, it asks about making review appointments while Questions 28 and 30 ask about the actual review appointment. Why are you focusing on the appointment scheduling? Are you actually asking about the times of day and days of the week for conducting the interview? Fourth, respondents can provide different types of answers for the phrases “times of day” (i.e., “mornings,” “two hours in the morning,” “8:00AM to 10:00AM”) and “days of the week” (i.e., “weekdays,” “midweek,” “Wednesday”). How do you want the respondent to answer?
- In Question 30, are you asking about the parts that are difficult for the reviewer, for the SNAP participant, or both? I can see a respondent being confused about what answer to provide.

- Two different time frames are used for Questions 31 and 31.A even though 31.A is based on 31. Furthermore, the answers will be confounded by time working as a QC reviewer. For example, Reviewer A was hired in 1985 and has coded increasingly more records as incomplete until 2006, when he has been coding fewer and fewer records as incomplete. However, the number of incomplete records is still greater now as compared to 1985. Therefore, he would respond that the number has increased since he has been conducting SNAP QC reviews. Reviewer B was hired in 2006 and has only seen the number of cases she has coded as incomplete decrease. Therefore, she would respond that the number has decreased since she has been conducting SNAP QC reviews. Both answers are valid but are they useful, especially since you are trying to compare State office procedures? Finally, the word “doing” seems a little too vulgar in this context. Could you change it to “conducting,” thereby removing any potential innuendo?
- Why are you asking for number of cases in Question 31? I ask this for two reasons. First, you asked for a percentage of cases in Question 21.C. Second, you stated that you wanted to avoid asking for direct numbers because the respondents had difficulty in providing hard numbers. Should you ask for the percentage?
- Question 32 is a test question and not a survey question. Change “the” before “main” to “your” to obtain the respondent’s answer.
- Why are you asking for number of cases in Question 32.A? I ask this for the same reasons listed for Question 31.
- What time frame are you using for Question 33?
- Question 34 has two issues. First, “general” is too ambiguous for the same reasons listed for Question 20. Second, this is another test question and not a survey question. Change “the” to “your” to get the respondent’s answer.
- What time frame are you using for Question 34.A?
- What does “typically” mean in Question 35? Does this mean for each case? Does this mean for cases that were originally assigned to the reviewer and no re-assigned? Does it mean each day? This term is ambiguous.
- What time frame are you using for Question 35.A?
- Question 36 contains multiple questions and should be separated. Furthermore, the first question is a test question and not a survey question. Change “the” to “your” to get the respondent’s answer.
- Question 37 contains multiple questions and should be separated. Furthermore, it seems that both questions in Question 38 are asking for yes/no answers. Do you want to know the incentives and penalties? If so, then asking direct questions about the incentives and penalties would be more effective.
- In Question 39, would the QC reviewers know who the regional office point of contact is for their States?
- Question 41 seems to ask for two different answers but leads the respondent into only providing one of those answers. The questions as about the feedback received for both complete and incomplete cases but emphasis is put on the feedback for the incomplete cases. Consequently, the respondent may feel that the answer for the incomplete cases is the only answer you want. What answers do you want? If you want information about the feedback for both complete and incomplete cases, then ask one question about the complete cases and another about the incomplete cases.

- Questions 41.A and 41.B seem to be asking for the same thing. What is the difference between the two questions?
- What is the purpose of the “Why or why not?” question in Question 44? I do not understand what you are trying to probe here.
- In Question 47.A, will the respondent know about the activities of the rest of the QC reviewers in that State?
- Does Question 49 include refresher training?
- Question 49.A contains multiple questions. However, questions two through seven seem to be probes and should not be included in Question 49.A. The interviewer should only ask them if the respondent did not provide them.
- What is the time frame for Question 49.B?
- Question 50 has four issues. First, it contains multiple questions and should be separated. Second, what is the time frame for the first question? Third, why is refresher training singled out? Fourth, what does the question “Annually?” mean? What do you want with it?
- What is the time frame for Question 50.A?
- Why are you asking question 53? It seems out of place since the previous questions have been focusing on overall training.
- Question 54 contains multiple questions and should be separated. Furthermore, it seems that the answer to the first question will be framed in terms of the training’s effectiveness. What if the training was ineffective? I would recommend adding the phrase “or ineffective” to the first question to remove any potential bias.
- In Question 55, who is the training helpful for? The respondent? Other QC reviewers? The State office? The question is not specific as to whom the training is helpful for.
- What do you want for Question 55.A? It seems that the respondent could give an answer that is technically correct but it is not what the respondent would do. Asking “What can you do to solve or minimize these problems?” will focus the answer on the respondent’s actions or recommendations.

State QC Reviewer Telephone Survey

- What time frame is the respondent comparing to in Question 5? Furthermore, the question should be reworded to be one question instead of two.
- In Question 6a, will the respondent know if the reviews are actually fairly distributed? This seems more like an opinion question. If a specific answer is needed and not just an opinion, then this question can be better answered by the supervisors or directors.
- The category answers for Question 6b are not mutually exclusive. A respondent could either legitimately offer multiple answers (e.g., “most months, but not all” and “more than once a year”) or ask for clarification on the answers. Furthermore, the terms “most,” “about,” and “a little more” are ambiguous and their meanings could vary among respondents. Finally, I think you mean “less than once a year.”
- The “all of the above” answer category in Question 7 is meaningless. Does it mean urban, rural, and suburban or does it include the two mix categories? If so, then why is there not a category for the urban/rural mix? I would recommend removing this option.
- How do you define “homeless” in Question 8? Will your respondents define it the same way?
- The answer categories are not distinct in Question 12. If it takes a reviewer 75 days to complete a QC review, does the reviewer select the second or the third category?
- The transition statement before Question 14 is too vague and it should be the same as the transition statement before Question 46.
- Remove “Would you say that you _?” from Question 14. It is unnecessary if you change the transition statement before this question.
- What number do you want the respondent to give in Question 18? I am assuming it is number of attempts.
- The answer categories are not distinct in Question 19.
- I do not know what kind of information you want in Question 20. In other words, what is the difference between the second and third option? What do those categories tell you? Furthermore, it seems that the respondent could answer yes to both the second and third categories.
- As for Question 20a, how do the answer categories relate to those in Question 20? For example, a respondent could report that 70-79% of the cases are completed after 2-3 contact attempts and less than 70% of the cases are completed after 4+ contacts? Which is the desired answer there? Furthermore, are these four categories enough to capture all of the data? In other words, if 95% of respondents selected the first option how would you use this data?
- A new time frame is introduced in Question 21: a review period. How long is a review period? Is it the same time lengths as listed in Question 12? Furthermore, are you asking the respondents to report for themselves or for the whole office? Adding “by you” after “QC review” will focus the question on the respondent.
- What is the time frame for Question 21a?
- Why do you need the first two answers in Question 22? It seems that they could be combined since the respondent is directed to do the same thing after answering either option.
- Question 23 is too ambiguous in that it could be asking for the number of contacts done by someone other than the respondent. Adding “for you” after “necessary” would focus the question on the respondent.

- Why are the formats of Question 24-29 different from the rest of the questions? You have previously been asking them direct questions whereas now you are asking them to verify statements.
- Why do you not have the respondents provide information if the answer “Other” for Questions 24b and 25a - 29a? While my question is not critical for the survey, it seems that you could gather some useful information this way.
- Question 31 is a test question and not a survey question. Change “the most effective” to “your most effective.”
- Question 32 is a test question and not a survey question. Change “the second most effective” to “your second most effective.”
- What is the time frame for Question 32a?
- In Question 33a, what do you mean by the word “doing”? It should be changed to “conducting” or something else to be more precise.
- What is the time frame for Question 35a?
- What kind of information do you hope to get by using the answer categories for Question 36? They are vague and you had used percentage categories in other questions.
- What time frame are you using for Question 37? Time is confounded in this question since the most recent incomplete case may be at very different times for different respondents. Furthermore, why is there not an “other” selection? Are the reasons listed the only reasons?
- Why do you not have the respondents provide information if the answer “Other” for Questions 38ai and 38aii? While my question is not critical for the survey, it seems that you could gather some useful information this way. Furthermore, why do you say “Some other reason” instead of the “Other” you have previously been using?
- Why do you not have the respondents provide information if the answer “Other” for Questions 41 and 42? While my question is not critical for the survey, it seems that you could gather some useful information this way.
- It seems that Question 43 should be part of a skip pattern with Question 42. If the respondent says “no” to any instructor-led training, why should he or she have to answer this question?

There are more issues with this questionnaire that I have not provided. Furthermore, I found many issues with the Supervisor/Director questionnaires that need to be fixed because they will affect the validity of your data.