Analysis of Measures of Respondent Burden
At the National Agricultural Statistics Service

Fatu Wesley
USDA, National Agricultural Statistics Service
1400 Independence Avenue, S.W., Washington, D.C., 20250

Keywords: burden, respondent, indicator, measure

The National Agricultural Statistics (NASS) uses joint burden indicators (JBIs) to measure the burden on respondents (especially those in the certainty strata). The JBIs are computed annually, by using information from each survey. The JBIs are 1) the total number of surveys a respondent is asked to complete during the year, 2) the total number of times a respondent is contacted and 3) the total number of minutes the respondent spends completing a survey. Analyses were performed to compare the average JBIs computed for seven surveys in order to 1) compare surveys in regard to the burden imposed on respondents, 2) determine the level of burden on respondents in the certainty strata versus respondents in other strata, and 3) find out which results from the analysis NASS can use to decrease respondent burden.

1. Indicators of Burden

The first indicator of burden is the number of surveys for which a farm operator is selected during the year. The second indicator of burden is the number of times that NASS contacts an operator during a year. The third indicator of burden is the total number of minutes estimated for a respondent to complete all surveys during the year. The estimate of the number of minutes is derived by NASS for each survey and is reported to the Office of Management and Budget (OMB). The number of surveys and the number of contacts will be the same unless surveys are coordinated so that the respondent completes more than one survey at a time. For this reason, the number of surveys is an upper bound for the number of contacts. The reason for using number of contacts as a burden indicator is that NASS's philosophy is that fewer contacts impose less burden. At this time, NASS has chosen to use three separate indicators to measure burden instead of a "burden index"; these indicators are referred to as the "joint burden indicators" (JBIs). Detailed descriptions of the burden indicators are given in *Use of Respondent Burden Measurement at NASS* (Scholetzky, 2005). Burden measurements are calculated for May 1 of one year to April 30 of the following year. This paper will focus on examination of the data generated from the burden measurements during May 1, 2005 to April 30, 2006.

2. Generating Respondent Burden Data

NASS Field Offices (FOs) administer national and their own state-wide surveys. NASS's Headquarters staff work with the FOs to ensure that polices for national surveys are standardized and managed correctly and efficiently across FOs. Procedures for improving and maintaining list frame sampling data is an example of this coordinated effort. NASS uses the Enhanced List Maintenance Operators database (ELMO) to manage its list sampling frame for the United States. Two of the tables in this relational database are the markedSample table and the surveyMaint table, which are used to compute the JBIs.

The time frame used to define "sampling year" in ELMO is "the annual cycle." This is the period in which the following occurs 1) control data is updated or added to the list sampling frame. Once the list frame is updated, a "frozen frame" is defined. 2) Populations are defined from the frozen frame for most surveys and 3) samples are selected for these surveys. The samples selected are stored in the markedSample table. The markedSample table contains the operators sampled for each survey. An entry in the markedSample table is uniquely identified by the *periodid* (used to identify the period in which the list frame is 'frozen' for sampling purposes), *stateid* (same as FO), surveyid (survey identification), *poid* (the operator id) and the *month*. If an operator is in a survey that is conducted twice a year, it will appear in the markedSample table twice (once for each month).

The second table, the surveyMaint table, is used by FOs, to record their survey coordination efforts. The primary fields in the surveyMaint table are *stateid, year, surveyid, periodid, month, surveylength, factor and coordCode*. *Surveylength* is the OMB number of minutes that is estimated for the operator to complete the questionnaire. *Factor* is the number of times the survey is conducted during the year. For example, suppose that a survey is conducted three times during the year. A FO can either set *factor*=1 and record data

on three lines of the surveyMaint table for each of the three months the survey is administered, or the FO can set *factor*=3 and record data for the survey on only one line of the surveyMaint table. The *coordCode* variable is used to identify surveys that are coordinated. A positive value of *coordCode* means that at least two surveys are coordinated. A value of zero means that the survey is not coordinated with another survey. All surveys that are coordinated are identified by the same value of *coordCode*. Each entry in the surveyMaint table must be linked to an entry in the markedSample table by *stateid*, *periodid*, *surveyid* and *month.*

An example of the surveyMaint table is in Table 1 and shows how a FO coordinated its surveys. Surveys 615,665 and 668 have a *coordCode* value of 1, which means that they were administered at the same time to operators that were in those surveys. The same holds for surveys 661,680 and 760, which have the same *coordCode* value of 2. Survey 940 has a *coordCode* value of zero, so this survey was not coordinated with any other survey.

Table 1: surveyMaint Values for a Field Office

Surveyid	Month	coordCode
615	2	1
661	12	2
665	1	1
668	1	1
680	1	2
940	11	0
760	1	2

Table 2 shows the actual burden data for a respondent. In the table, *factor* is the number of times the survey is administered. Survey 182 is administered twice during the year, so it has a value of 2. The month code for Survey 182 is 6, which is the first time the survey is conducted. Survey 661 and survey 187 are coordinated (both have coordCode=12), so this reduces the number of contacts for this respondent by 1, which makes a total of six contacts for the year.

Note that although the *coordCode* values for surveys other than 182 and 661(month12) are positive, only these two surveys are coordinated for this operator. The number of minutes is 160, which is the sum of the total OMB time for the seven surveys combined. Survey 182 has a factor of 2, so its total *surveylength* is 90 (45x2) minutes.

Table 2: An Example of Burden Data for a Respondent

Survey	Month	Factor	coordCode	Survey Length	Number of Surveys	Number of Contacts	Number Of Minutes
661	12	1	12	20	7	6	160
661	3	1	3	20	7	6	160
190	5	1	5	10	7	6	160
182	6	2	6	45	7	6	160
187	10	1	12	15	7	6	160
413	4	1	4	5	7	6	160

3. Scope of the Analysis

The seven surveys that were used in the analysis were Agricultural Resource Management (ARMS), Agricultural Yield (AgYld), Quarterly Agricultural Survey (QAS), Cattle, Hogs, Agricultural Labor (Labor), and Sheep/Goats. There were 13 FOs included in the analysis. The strata compared were certainty and non-certainty strata. A certainty stratum is one in which all sampling units have a probability of selection equal to 1. The data used to compare surveys was the average value of each indicator for all operators selected for a survey within a FO. The data used to compare strata was the average value of each indicator for all operators assigned to a certainty/non-certainty stratum (within a FO and survey). For example, a value used in the analysis comparing surveys was the average for operators selected in Field Office x for survey y. A value used in the analysis comparing strata was the average for operators selected in Field Office x for survey y and stratum z. For these reasons, the rest of this paper will refer to the 'average' of the burden indicators.

4. How Surveys Compared

The results based on comparing surveys have to be interpreted with the understanding that the burden indicators are assigned to the respondent, not the surveys (see Table 2). The burden indicators assigned to an operator do not vary by the survey for which the operator is selected - it is the same for all the surveys for which the respondent is selected. If an operator is selected for the Cattle, Hogs, Labor and Sheep/Goats surveys, its number of surveys is four. If the interview for cattle and hogs are coordinated, its number of contacts is three. The respondent's number of minutes is the sum of the number of minutes needed to complete the four surveys. Note that surveys that are *not* included in the analysis affect the averages if they are in the surveyMaint table and the operator is selected for those surveys.

In spite of the way in which the data is recorded, surveys can be compared because the more burdensome a survey is (number of times it is conducted, etc.), the more it appears in the surveyMaint table, and the more its characteristics influence the level of the burden indicators.

Table 3 gives the mean indicator value by survey. Keeping in mind that the ARMS survey has *factor*=2 (because it is conducted twice during the year), the interpretation of the averages can be illustrated using the ARMS values:

"Operators selected for the ARMS survey are selected for an average of four surveys a year, *including* ARMS."

"Operators selected for the ARMS survey are contacted an average of four times a year, *including* ARMS."

"Operators selected for the ARMS survey are interviewed an average of 118 minutes a year, *including* ARMS."

Table 3. Average Value of Burden Indicators by Survey			
Survey	Average Number of Surveys	Average Number of Contacts	Average Number of Minutes
ARMS	4 .13	4.04	118.06
AgYld	6.50	6.34	88.57
QAS	4.94	4.57	89.43
Cattle	4.08	3.88	73.12
Hogs	4.18	3.66	56.82
Ag Labor	4.37	4.27	83.71
Sheep & Goats	2.25	2.17	43.03

A test for normality showed that the data were not normally distributed, because of the small sample size (13 FOs), so the Kruskal-Wallis test (nonparametric alternative to the one-way analysis of variance test) was used to compare the surveys.

The Kruskal-Wallis is computed by assigning a rank to each observation in the analysis. An observation is one combination of a FO and survey. Thus the 91 observations (7 surveys x 13 FOs) are ranked from 1 to 91. The average of the ranks is then computed by survey. The average ranks are used to compute the test statistic. The hypotheses tested for each burden indicator were:

H_o: $M_{ARMS} = M_{AgYld} = M_{QAS} = M_{Cattle} = M_{Hogs} = M_{Labor} = M_{Sheep/Goats}$

H_1: $M_{ARMS} \neq M_{AgYld} \neq M_{QAS} \neq M_{Cattle} \neq M_{Hogs} \neq M_{Labor} \neq M_{Sheep/Goats}$

Where M=mean rank

The means were computed using Proc Npar1way in SAS. The p-values of the Kruskal-Wallis tests show that the differences among the surveys are significant for all indicators. The values are:

Average Number of Surveys:
(χ^2 =58.63, df=6, p-value < .0001)

Average Number of Contacts:
(χ^2 =58.02, df=6, p-value < .0001)

Average Number of Minutes:
(χ^2 =68.21, df=6, p-value < .0001)

The results of the Kruskal-Wallis tests show that for each burden indicator, at least two surveys are significantly different; they do not tell us which ones are different. When a statistical test shows that more than two items are significantly different, the follow-up analysis is a multiple comparison test to show which items are different. The nonparametric pairwise multiple comparison procedure given in Conover (1999) was conducted.

The results of the tests for average number of surveys showed that the only surveys that are different *from all other surveys* were QAS, AgYld and Sheep/Goats. Farm operators selected for the AgYld survey were selected for about 7 surveys; QAS about 5, and Sheep/Goats operators are selected for about 2 surveys. The results of average number of contacts were similar to the results for average number of surveys. Again, QAS, AgYld and Sheep/Goats were unlike the other surveys and each other. AgYld had the highest number of contacts at about 6, and Sheep/Goats has the lowest number at about 2.

The results for the average number of minutes show that the only surveys that are alike are QAS & AgYld and Cattle & Labor. Sheep/Goats operators had the lowest number of minutes and ARMS had the highest. Operators that are selected for the Sheep/Goats survey spend about 43 minutes completing questionnaires for *all* surveys for which they are selected, and operators selected for the ARMS survey spend about 118 minutes completing questionnaires for *all* surveys for which *they* are selected.

5. How Strata Compared

The comparison of strata includes only four surveys: Cattle, QAS, Hogs and Sheep/Goats. ARMS and AgYld do not have certainty strata. The Labor survey has certainty strata but they exist in only four FO's, so Labor was excluded to simplify the analysis. The comparisons were conducted within survey because an operator that is in the certainty strata of one survey may not be in the certainty strata of another survey. For example, an operator that is in both the Labor and Hogs surveys may be in the certainty strata of the Labor survey because it hires a lot of workers, but not in the certainty strata of the Hogs survey, because it is not a large hog operator.

A two-sample nonparametric test (Kruskal-Wallis) was used to determine if the differences between strata were significant. The hypotheses tested here for each indicator and all surveys were:

H_o: $\mu_{Survey\ i,\ Certainty} = \mu_{Survey\ i,\ Non\text{-}Certainty}$
H_1: $\mu_{Survey\ i,\ Certainty} \neq \mu_{Survey\ i,\ Non\text{-}Certainty}$

Where i = Survey (Hogs or Cattle, etc.)

Mean differences between the certainty and non-certainty strata were significant for all surveys and burden indicators. The Cattle survey had the biggest difference between the two types of strata for number of surveys, number of contacts and number of Minutes.

Operators in the Cattle sample who are in the certainty strata were selected for about five samples more than operators in the cattle sample who are in the non-certainty strata. Operators in the Cattle sample and in the certainty strata were contacted about 4.7 times more than operators in the Cattle sample and in the non-certainty strata. The Sheep/Goats survey had the smallest differences between the certainty and non-certainty strata. Operators who were selected and contacted for the Sheep/Goats survey were selected and contacted about 1.5 times more when they are in the certainty strata.

Average time differences between the certainty and non-certainty strata were at least 30 minutes for all surveys. Operators in the Cattle sample and the certainty strata spend about 90 additional minutes completing surveys than operators in the Cattle sample and the noncertainty strata, while operators in the Sheep/Goats sample and the certainty strata spend about 30 additional minutes completing surveys than operators in the Sheep/Goats survey and the noncertainty strata.

6. Are we Reducing Respondent Burden?

Although some respondents might prefer two contacts to complete two questionnaires rather than one contact where a longer time is spent completing the same questionnaires, the difference between the number of surveys and the number of contacts is still a good test to determine the extent to which respondent burden is being reduced. The difference between the two indicators was computed (for each operator) and a nonparametric paired test was used to determine if the difference was significant. *The analysis was restricted to operators in the certainty strata within the same survey*.

The hypotheses tested were:

H_o: $\mu_{\text{Difference}} = 0$
H_1: $\mu_{\text{Difference}} \neq 0$

The Wilcoxon sign test was conducted to determine if the paired differences between the average number of surveys and average number of contacts were significant. Results presented in Table 4 show that the differences between number of surveys and number of contacts are significant for all surveys. Given that the mean difference for all surveys is less than one, it is not clear if statistical significance translates into practical significance.

Table 4. **Certainty Strata: Differences between Number of Surveys and Number of Contacts by Survey using the Wilcoxon Signed Test**

Survey	Number of Surveys	Number of Contacts	Difference	p-value
Cattle	8.59	8.00	.59	.0001
QAS	7.59	6.66	.93	.0001
Hogs	6.73	6.02	.71	.0001
Sheep/ Goats	3.71	3.47	.24	.0001

7. Conclusions

An unexpected finding of the respondent burden analysis is that of the 49 field offices that were instructed to coordinate surveys to reduce respondent burden, only 14 (29 percent) did so. One Field Office showed coordination of their state-level surveys only, and was excluded from the analysis. The analysis was based on data for the 13 FOs that coordinated national surveys.

Comparisons of the data for 13 FOs using the nonparametric Kruskal-Wallis tests showed significant differences among surveys for each of the burden indicators. Pairwise comparison procedures performed to determine which surveys were different showed that for number of surveys and number of contacts, operators selected for the AgYld survey had the most burden and operators selected for the Sheep/Goats survey had the least burden. Concerning time spent completing the questionnaires, respondents that were selected for the ARMS survey had the most burden, and those for the Sheep/Goats survey had the least burden.

The data for the 13 FOs were also compared by type of strata. The types of strata were certainty and non-certainty. The differences between the certainty and non-certainty strata were significant for each survey and burden indicator. The Cattle survey had the largest differences between strata and the Sheep/Goats survey had the smallest differences.

Within the certainty strata, the number of surveys and the number of contacts were compared to determine how much burden was reduced. The comparisons were done within survey. The results indicate that the differences between number of surveys and number of contacts were significant. Although the results were significant, the average difference between the number of surveys and the number of contacts was less than one.

8. Recommendations

1 Field Offices that coordinate surveys should maintain the surveyMaint table. Coordinating surveys *and* maintaining the surveyMaint table should be one of the measures by which Field Offices are rated annually.

2. Once Field Offices start maintaining the surveyMaint table, they should be given annual feedback about how successfully they are reducing burden. These results should be published on the NASS intranet.

3. This analysis should be considered preliminary. After all FOs are compliant with maintaining the surveyMaint table, a more complete analysis should be conducted. This will also create a larger data set so that more powerful parametric tests can be used.

4. Although differences between the number of surveys and number of contacts were significant, given that *they were all less than one*, there seems to be room for improvement in our efforts to coordinate surveys.

5. The surveys that are the most burdensome, in terms of time needed for completion, should be examined to determine if they can be modified to reduce the amount of time required to complete them.

6. Asking respondents how we can reduce their burden might be useful. A random selection of respondents in AgYld and ARMS surveys could report their views about number of surveys versus number of contacts - do they prefer two short contacts to one longer contact to complete the questionnaires for two surveys?

7. If possible, the actual number of minutes to complete a questionnaire should replace the OMB number of minutes in the surveyMaint table. This will provide more realistic information.

8. The issue of computing a burden index should be revisited. It might be possible to use techniques such a

factor analysis to create a usable index. A burden index would be a useful tool for Field Offices.

References

1. Scholetzky, Wendy (2005), "Use of Respondent Burden Measurements at NASS", in *Proceedings of the Section on Government Statistics*, American Statistical Association, pages 1745-1749

2. Respondent Burden Tracking Team (2000), "The Respondent Burden Tracking and Management System", USDA/National Agricultural Statistics Service.

3. Conover, W.J.(1999), *Practical Nonparametric Statistics*, 3rd Edition, John Wiley & Sons, Inc.