

**And the Survey Data Says:  
Top Facts about Our Online Respondents and  
How We Plan to Get More of Them**

Jaki McCarthy and Chris Gottschall<sup>1</sup>

**Abstract**

The 2007 Census of Agriculture was the first to offer online reporting and nearly 100,000 Census of Agriculture reports were received online. We examined traits of web respondents and quality of their data. Web respondents were smaller, newer, more urban farms with high speed internet access. Their data were also better quality. Thus, several experiments were designed to target and increase web response during the testing for the 2012 Census of Agriculture. Comparisons were made between: groups 1) mailed a cover letter and instructions for online reporting, without a paper form; and 2) mailed a cover letter requesting them to report online but also included a paper form. In addition, several different cover letter treatment groups were tested. Response rates (both online and overall) and plans for increasing online reporting for the 2012 Census of Agriculture will be discussed.

## **1. INTRODUCTION**

The 2007 Census of Agriculture (COA) was the first installment of the COA to utilize a web reporting option. Prior to this, many National Agricultural Statistic Service (NASS) surveys had successfully incorporated an internet reporting feature into the agency's multiple modes of data collection, but the 2007 COA was by far NASS's largest online reporting effort to date. For the COA, paper forms were mailed to all known or potential agricultural operations. Each form also included a URL where a respondent could provide their data instead of completing and mailing back the paper form. Aside from providing login information in the mailings, no special attempts were made in the 2007 COA to encourage online reporting. However, nearly 100,000 reporters completed the online COA. Web response rates varied substantially between states. Using data mining techniques, we were able to determine what influences respondents to report via the web. We also analyzed the data collected in online reports in order to evaluate characteristics of these farm operations, completeness of their reports, and overall data quality. We then used this information to design several experiments that would capitalize on what we had learned about our online respondents in an effort to increase our web response rates.

The purpose of this project was multi-faceted: 1) Analyze the data results and respondent characteristics of the 2007 Census of Agriculture; 2) Draw conclusions from the results and develop possible ways to increase our web response rates; 3) Test the options, and based on the results, arrive at recommendations for the 2012 Census of Agriculture.

---

<sup>1</sup> Authors are with the USDA, National Agricultural Statistics Service Research and Development Division and Data Collection Branch, respectively. Special thanks to Thomas Jacob (USDA, NASS, Information Management Group) whose extensive analysis of the 2007 COA results provided the framework for this experiment.

## 1.1 2007 Census of Agriculture Analysis Overview

The 2007 COA results were analyzed using advanced data mining techniques. Comparisons were made between mail and web respondents. However, it is important to note that this was not an experimental comparison of records assigned to each mode of response, but rather an analysis of which respondents chose each mode. Both mail and web reporting were available to all respondents. This allowed us to hone in on the most important traits and differences of the web and mail respondents. The analysis consisted of three major parts: an initial review of the list/control data available for the records prior to the mail-out of the COA; a summarization of the reported data for both groups in order to find differences as well as any indicators of the propensity to respond via the web; and an assessment of the overall completeness and quality of the reported data from the two modes.

## 1.2 2007 COA List Data Analysis

Prior to the 2007 COA, extensive list building and cleaning efforts took place. List maintenance is typically an ongoing, dynamic process within NASS; however, preparation of the Census Mail List (CML) necessarily led to additional review and scrutiny, over and above regular, daily operations. Records were assigned an active status code as part of the list building and maintenance process. In general, these codes are based on when a record was added, its past contact history, and what is contained in the list about that particular record.

**Figure 1:** Active Status Codes of Mail and Web responses

Active Status Code	MAIL (%)	WEB (%)
<b>0 – Active Farm Operation</b>	<b>74.60</b>	<b>68.20</b>
9 – Special COA use only	0.40	0.10
<b>30 – Screening, Potential Farm</b>	<b>8.60</b>	<b>13.40</b>
31 – Screening, Previously inactive	1.60	1.30
32 – Screening, CRP presence	3.40	3.40
33 – Screening, Potential future sales	0.60	0.80
<b>34 – Screening, Nonrespondent</b>	<b>8.80</b>	<b>10.20</b>
35 – Screening, Refusal	0.20	0.20
<b>36 – Screening, Specialty</b>	<b>1.70</b>	<b>2.40</b>
<b>Total</b>	<b>100.00%</b>	<b>100.00%</b>

The key thing to notice in this analysis is that Active Status 30, 34, and 36 are all new additions to the list that have not been included in any of the agency’s surveys except a short screening survey intended to determine whether the respondents qualify as a farm by USDA definition. More web respondents tended to be recent additions to NASS’s list frame. These could either be new operations or farms that had been missed in previous list building efforts. The mean “age” of the records (when they were originally added to the list) was also calculated. Web respondents tended to be substantially newer additions than mail respondents.

One year prior to the 2007 COA, a screening survey was conducted to evaluate the recent additions to the list, operations whose status was questionable, etc. and to gather limited control data for the operations prior to adding them to the CML. Twenty percent of the web respondents were added to the CML based on the screening survey results, compared

to 17 percent of the mail respondents. Additionally, of those records that responded via the web, only 18 percent had been contacted for another NASS survey in the last three years. Since NASS survey samples are based on control data, this also lends support to the emerging trend of web respondents being newer additions with limited control data. The 2007 COA consisted of two versions: a long form for those on the CML that were, based on existing information, expected to report multiple commodities or complex agriculture operations; and a short form for less complex operations, smaller operations and newer operations for which NASS had little known information. No matter which mail form group a record was assigned to, the respondent was given the option of reporting via the web or returning the mail form.

**Figure 2:** Assignment to Long or Short Form mailing groups for Mail and Web responses

<b>Assigned Form</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
Long Form	81.0	85.1
Short Form	19.0	14.9
<b>Total</b>	<b>100.00%</b>	<b>100.00%</b>

This was somewhat counter to the previous findings: based on the newness of the typical web respondent’s record, available list data, and sampling/reporting history, it was expected that a larger proportion of the web respondents would have received the short form. Upon further analysis of the reported data, it did appear that very large farms (which would have been assigned to the long form mailing group) also had a tendency to report via the web, possibly explaining this difference.

Even with these list building and cleaning efforts, there were still records on the CML that ultimately screened out, did not meet the NASS farm definition, were out of business, never a farm, etc. These records are considered “out of scope” for purposes of the COA.

**Figure 3:** Current Status Codes of Mail and Web responses

<b>Final Current Status Code</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
4 – Assigned for followup	0.30	0.00
23 – Undeliverable As Addressed	0.20	0.00
33 – Refusal	0.20	0.20
51 – Duplicate	1.60	0.50
52 – Deceased	1.60	0.50
53 – Landlord only	0.30	0.10
<b>54 – Out of Scope: Non ag, never farmed</b>	<b>1.20</b>	<b>13.10</b>
56 – Retired or disabled	0.20	0.00
<b>57 – Out of Scope: Other</b>	<b>9.10</b>	<b>12.80</b>
<b>58 – Computer Classified Out of Scope</b>	<b>21.80</b>	<b>11.00</b>
59 – Reviewed and Out of Scope	0.80	0.60
<b>61 – In Scope</b>	<b>62.70</b>	<b>61.20</b>
<b>Total</b>	<b>100.00%</b>	<b>100.00%</b>

The proportion of current status code 61 – In Scope records was almost the same for both mail and web respondents. However, the proportions of several of the other current status codes differed between the two response types. The web reporting instrument took

respondents through a series of screening questions prior to entering the survey proper. Screening out assigned the record to a non-ag status, which explains the increased number of web responses assigned to out of scope status code 54. There are no equivalent screening questions on the paper form, so a determination that the record is out of scope is made from the data that is entered on the form. (This might be, for example, that the respondent entered 0 acres operated.) The “computer out of scope” records (status code 58) were records that the edit determined not to be farms. Note that for the mail returns, these would have to have been processed and key entered—a fairly expensive, time consuming step. From these results, it appears that the web reporting system/instrument may be a more efficient way to capture records that are out of scope before they are subjected to additional (and much more expensive) processing and handling.

Following this initial meta-analysis of the respondents, some more record level, control type data review was conducted. Consistent with the limited control data, 42.5 percent of the web respondents did not have a control data calculated farm type assigned, compared to 32.2 percent of the mail respondents. Based on existing list data, mail respondents were more than 10 percent older than web respondents. Three times the web respondents had an email address already stored in the list data. Using externally available zip code population density data, web respondents tended to live in much more densely populated areas: 208 people per square mile for web respondents versus 161 people per square mile for mail respondents. Mail and web respondents were also compared based on their Rural-Urban Continuum code<sup>2</sup>. This coding scheme categorizes each US county based on its population and distance to metropolitan areas. Using this measure, web respondents also appear much more urban than mail respondents as shown below.

**Figure 4:** 2003 Rural-Urban Continuum Codes for Mail and Web responses

<b>2003 Rural-Urban Continuum Codes</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
1 (most urban)	<b>15.20</b>	<b>19.10</b>
2	<b>13.90</b>	<b>15.70</b>
3	<b>13.00</b>	<b>13.20</b>
4	9.00	8.80
5	3.60	3.60
6	21.20	18.70
7	11.70	10.40
8	5.10	4.20
9 (most rural)	7.30	6.30
<b>Total</b>	<b>100.00%</b>	<b>100.00%</b>

### 1.3 2007 COA Reported Data Analysis

Analyzing the pre-existing list/control data for mail and web respondents was the first step in the analysis. The second step was to examine the actual reported data for differences between the two groups. For these comparisons, the analysis was conducted just on those COA records determined to be in scope.

<sup>2</sup> The Rural Urban Continuum codes and their descriptions are available at <http://www.ers.usda.gov/Data/RuralUrbanContinuumCodes/>.

**Figure 5:** High Speed Internet Access for Mail and Web responses

<b>Have High Speed Internet Access?</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
Yes	31.0	66.8
No	69.0	33.2
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

The internet access results were consistent with our original hypothesis that the people with internet access in general and high speed service in particular, would be more likely to respond via the web. However, this may also be more indicative of an underlying trait: respondents are more open to using new technology and would obtain high speed internet access as it is available, whereas others may be more ambivalent towards technology, and thus less likely to acquire better access and use the internet for business. It is interesting to note that a third of the web respondents did not report having high speed access, and were still willing to complete the COA online.

This is also consistent with our hypothesized web respondent profile: small/hobby or lifestyle farmers, closer to urban areas (and hence better/faster internet access options), deriving more of their income from off the farm, etc. These respondents would generally be more familiar or comfortable using computers and the internet in their daily lives and for conducting business, submitting information, and/or contacting different public entities.

As noted in the previous section, the Census mailing list consisted of both short and long form groups. Both of these were given the option of reporting on the web. However, the web instrument was the same for both short and long groups. The design of the instrument required the respondents to actually “page” through the entire web form which was very similar to the long form. This resulted in more data and more complete forms compared to the mail returns. These results are also consistent with previous research on effective design for collecting data from web respondents (Dillman 2000; Couper, Traugott, and Lamias 2001; Couper 2000; etc). Given the design of the web instrument, we expected this from the outset, but the magnitudes were somewhat more substantial than originally hypothesized.

**Figure 6:** Mean number of data items for Mail and Web respondents, by mailing group

<b>Original Mailing Group</b>	<b>Mean number of data items (per report)</b>	
	<b>MAIL</b>	<b>WEB</b>
Long Form	89.1	96.0
Short Form	63.8	83.4

The records were all assigned to the long or short form mailing groups based on the a priori expected amount of agricultural activity, e.g. small hobby/lifestyle farmers generally received the short form. These types of operations would typically be expected to have fewer data items to report, resulting in something like the results of the mail respondents. Obviously, some of the difference in the number of items for the short form respondents is explained by the fact that the web respondents completed essentially a long form online. However, this does not explain the difference in the number of items for both mail and web respondents that were in the long form group. Additionally, note

that the number of data items reported is comparable between the two mailing groups for the web respondents, whereas there is a substantial difference in the number of data items for mail respondents depending on the form completed.

Taking this analysis in a slightly different direction, the number of data items reported was analyzed by farm type for mail and web respondents. Web respondents consistently had more data items, regardless of farm type.

**Figure 7:** Mean number of data items for Mail and Web respondents, by calculated farm type

Calculated Farm Type (based on reported data)		Mean number of data items (per report)	
		MAIL	WEB
1	Grains, Oilseeds, Dry Beans, and Dry Peas	94.5	103.3
2	Tobacco	91.6	101.4
3	Cotton	95.0	103.3
<b>4</b>	<b>Vegetables, Melons, Potatoes, and Sweet Potatoes</b>	<b>94.3</b>	<b>110.1</b>
5	Fruit, Tree Nuts, and Berries	84.2	94.2
6	Nursery, Greenhouse, Floriculture, and Sod	87.4	95.0
7	Cut Christmas Trees and Short Rotation Woody Crops	80.5	85.6
<b>8</b>	<b>Other Crops and Hay</b>	<b>67.9</b>	<b>83.2</b>
9	Hogs and Pigs	96.6	102.9
10	Milk and other dairy products from cows	107.9	115.6
11	Cattle and Calves	87.5	95.5
12	Sheep and Goat	84.6	90.6

Key to note is that especially for the farm types consistent with the developing “profile” of who the web respondents were, there were many more data items collected. This further suggests what was previously noted: requiring respondents to work through the entire instrument tends to result in more data items (Dillman 2000). (Of course, the alternative explanation that web respondents simply had more items to report is also possible.) This provided additional insight into the web instrument’s ability to collect data.

The total value of production is calculated for respondents as part of the COA processing based on their reported data. Proportions of web respondents were higher on both ends of the spectrum. The smallest farms (typically hobby or “lifestyle” farms) and the largest operations both tended to report via the web.

**Figure 8:** Total Value of Production for Mail and Web respondents

<b>Total Value of Production Ranking</b>		<b>MAIL</b> (%)	<b>WEB</b> (%)
P	No sales	19.6	18.2
1	\$1 – 999	5.7	7.7
2	\$1,000 – 2,499	9.0	10.6
3	\$2,500 – 4,999	9.4	10.1
4	\$5,000 – 9,999	10.6	10.2
5	\$10,000 – 24,999	13.1	11.4
6	\$25,000 – 49,999	7.8	6.7
7	\$50,000 – 99,999	6.5	5.3
8	\$100,000 – 249,999	7.6	6.6
9	\$250,000 – 499,999	4.8	5.1
10	\$500,000 – 999,999	3.2	3.9
11	\$1,000,000 – 2,499,999	1.8	2.6
12	\$2,500,000 – 4,999,999	0.4	0.6
13	\$5,000,000 or more	0.2	0.6
<b>Total</b>		<b>100.0%</b>	<b>100.0%</b>

Rounding out the data analysis for mail and web respondents was an assessment of the demographic items. Web respondents tended to work off the farm, spend more days working off the farm, and have a higher household income. Slightly more web respondents reported female primary operators.

**Figure 9:** Primary Occupation for Mail and Web respondents

<b>Primary Occupation</b>	<b>MAIL</b> (%)	<b>WEB</b> (%)
1-Farming	47.1	38.7
2-Other than farming	52.9	61.3
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

**Figure 10:** Number of days worked off the farm for Mail and Web respondents

<b>Number of Days Worked Off Farm</b>	<b>MAIL</b> (%)	<b>WEB</b> (%)
None	36.4	26.2
1-49 days	11.2	8.9
50-99 days	5.3	4.8
100-199 days	9.0	9.5
200 days or more	38.0	50.6
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

**Figure 11:** Household Income for Mail and Web respondents

<b>Household Income</b>	<b>MAIL</b> (%)	<b>WEB</b> (%)
Less than \$20,000	14.7	9.8
\$20,000 – 29,999	9.7	6.5
\$30,000 – 39,999	13.4	9.1
\$40,000 – 49,999	12.2	10.8
\$50,000 or more	50.0	63.8
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

**Figure 12:** Gender of the primary operator for Mail and Web respondents

<b>Gender of the Primary Operator</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
Male	87.6	85.4
Female	12.4	14.6
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

#### 1.4 2007 COA Data Quality Analysis

The final phase of the 2007 COA analysis was an evaluation of the overall quality of the in-scope, reported data for web respondents compared to the mail respondents' data. This was measured three ways: the number of analyst interventions in order to clean the record; the average number of imputed items per record (a measurement of item missing data); and the proportion of records with no item imputation at all.

**Figure 13:** Analyst interventions for Mail and Web respondents

<b>Analyst Interventions</b>	<b>MAIL (%)</b>	<b>WEB (%)</b>
Clean as reported	54.4	56.5
Edited twice	19.6	20.0
Clean, but reviewed	14.6	11.5
All Others	11.4	12.0
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

The web responses tended to be cleaner as reported. The margin increased as the records that were edited twice (reported data plus one analyst edit) were included. Additionally, the number of records that were clean but reviewed was lower for web reports. This is most likely an indication of a more consistent dataset, with fewer outliers requiring analyst review.

Consistent with these results, is the number of COA records that required no item level imputation: 21 percent of the web responses required no item imputation, whereas only 12 percent of the mail returns contained complete data for which no item level imputation was necessary. Of the records that did require item level imputation, mail responses required more imputed items in order to be considered complete.

**Figure 14:** Number of imputed items for Mail and Web respondents

	<b>MAIL (mean)</b>	<b>WEB (mean)</b>
<b>Average number of imputed items per record</b>	6.41	4.18

## 2. METHODS

From the results of the 2007 Census of Agriculture analysis, web respondents tended to be from smaller, newer, more urban lifestyle farms with ready access to high speed internet service. Their data also tended to be somewhat more complete and "cleaner" as reported. Armed with this knowledge (and visions of increased efficiency and cost savings), we constructed several experiments designed to target these types of operations

and encourage them to respond via the web during the testing preceding the 2012 Census of Agriculture. We were able to develop some good ideas on possible ways to target respondents in order to encourage a web response. These were generally similar to previous research (Millar, O’Neill and Dillman 2009) with a key exception: we had fairly rich list data (coupled with reported data from the COA) allowing us to hone in and target certain respondents with these determining attributes.

From the analysis of the reported data, access to high speed internet service was the primary determinate of the propensity to respond via the web. We drew a sample of the respondents who had all reported having high speed internet access in 2007 COA. We included both those who had responded to NASS via the web previously and records with high speed internet access that had not responded via the web. Our primary objective was to get those records with high speed internet access to respond via the web. The Census of Agriculture will most likely always be a multi-mode data collection effort. It is simply the most efficient way to collect the required volume of data in a timely manner. However, given the results of the data analysis, pushing those respondents most likely to respond via the web into actually reporting online could result in increased efficiencies through cheaper, faster, better and more complete data.

The sample was split between two different treatments: 1) A “Web” group which was mailed a cover letter and instructions on completing their Census of Agriculture via the web, but were NOT mailed a paper form; and 2) A “Web/Paper” group whose mailing also included a paper copy of the Census of Agriculture as an alternative to the recommended web reporting. All records in both treatment groups had previously reported having high speed internet access. The control groups for each treatment were records who had reported high speed internet access and had responded via the web. Within each primary treatment group, there were also three experimental subgroups: 1) Stating in the cover material that we were asking respondents with high speed internet access to report via the web; 2) Presenting a due date after which respondents would receive follow-up mailings or phone calls; and 3) A combination of both the high speed internet access mention and a due date.

**Figure 15:** Summary of experimental treatment groups

<b>Treatment Group</b>	<b>Initial Mailing</b>	<b>Additional Test</b>
Web A	“Special” web reporting letter only	Due date
Web B	“Special” web reporting letter only	Due date and asked those with high speed internet access to respond online.
Web C	“Special” web reporting letter only	None
Web D	“Special” web reporting letter only	Asked those with high speed internet access to respond online.

*(Continued next page)*

<b>Treatment Group</b>	<b>Initial Mailing</b>	<b>Additional Test</b>
Web/Paper A	Questionnaire and “special” web reporting letter	Due date
Web/Paper B	Questionnaire and “special” web reporting letter	Due date and asked those with high speed internet access to respond online.
Web/Paper C	Questionnaire and “special” web reporting letter	None
Web/Paper D	Questionnaire and “special” web reporting letter	Asked those with high speed internet access to respond online.
Main Group	Questionnaire with stock cover letter (contains web reporting instructions)	None

For this particular test, the treatment group mailing only applied to the initial mailing package. In the subsequent follow-up mailings to initial non-respondents, all received the same mailing package (a stock cover letter, web reporting instructions and a paper questionnaire).

### **3. RESULTS**

In early 2010, the Census of Agriculture content pretest was conducted. The final sample size was just over 4,800. Of these, approximately 40 percent had previously reported having high speed internet access and were selected for the web response test. Early on, we experienced some logistical and handling challenges as well as some mail-out delays. This ultimately affected the number of records available for the analysis. There were also some inconsistencies in the mailing packages which we speculate may have led to some confusion on the part of the respondents. Overall, there were enough data for meaningful analysis, but due to some of the challenges we faced, we plan to conduct at least one more round of similar experiments prior to finalizing any plans for the 2012 Census of Agriculture.

#### **3.1 Web Groups**

The Web group was the portion of the sample that had previously reported having high speed internet access and was not mailed a paper questionnaire in the initial mailing in an effort to “force” them to respond via the web. There were four different treatment groups within the sample (see Figure 14: Summary of experimental treatment groups, page 9), all of which contained respondents who had previously reported via the web and those who had not previously reported via the web.

**Figure 16:** Web groups, web response, mail response and non-response rates by previous web reporting history

Treatment Group	Web Response (%)	Mail Response (%)	Non-response (%)
<b>Web A</b>			
Previous Web Response	30	0	70
No Previous Web Response	23	0.5	76.5
<b>Total</b>	<b>25</b>	<b>0.4</b>	<b>74.6</b>
<b>Web B</b>			
Previous Web Response	29	0	71
No Previous Web Response	28	0.5	71.5
<b>Total</b>	<b>28</b>	<b>0.4</b>	<b>71.6</b>
<b>Web C</b>			
Previous Web Response	20	2	78
No Previous Web Response	15	1	84
<b>Total</b>	<b>16</b>	<b>1.5</b>	<b>82.5</b>
<b>Web D</b>			
Previous Web Response	21	0	79
No Previous Web Response	18	0	82
<b>Total</b>	<b>19</b>	<b>0</b>	<b>81</b>
<b>Overall Web Groups</b>	<b>22</b>	<b>0</b>	<b>78</b>
<b>Main Group (control)</b>	<b>5.0</b>	<b>28.6</b>	<b>66.4</b>

These results were somewhat mixed and not exactly as expected. The Web C and Web D treatment groups experienced some handling problems, possibly explaining the decreased web response rates. It was encouraging to see that when mailings and logistics were executed correctly (in the Web A and Web B treatment groups), we were able to indeed force a larger portion of respondents to the web—typical NASS survey web response rates are around 3-5 percent. However, this increase came at the expense of the overall response rate. In the second mailing, the Web group non-respondents were also mailed a paper questionnaire and given the option of returning it or responding via the web. Mail returns were virtually non-existent, especially when compared to mail response rates of the Web/Paper groups (next section).

### 3.2 Web/Paper Groups

The Web/Paper group was the portion of the sample that had previously reported having high speed internet access and was mailed a paper questionnaire in the initial mailing but was encouraged to report via the web in an effort to “push” them towards online responding. Much like the Web group, there were four different treatment groups within the sample (see Figure 14: Summary of experimental treatment groups, page 9), all of which contained respondents who had previously reported via the web and those who hadn’t previously reported via the web.

**Figure 17:** Web/Paper groups, web response, mail response and non-response rates by previous web reporting history

Treatment Group	Web Response (%)	Mail Response (%)	Non-response (%)
<b>Web/Paper A</b>			
Previous Web Response	34	9	57
No Previous Web Response	24	13	63
<b>Total</b>	<b>26</b>	<b>12</b>	<b>62</b>
<b>Web/Paper B</b>			
Previous Web Response	33	7	60
No Previous Web Response	27	13	60
<b>Total</b>	<b>29</b>	<b>11</b>	<b>60</b>
<b>Web/Paper C</b>			
Previous Web Response	32	10	58
No Previous Web Response	25	12	63
<b>Total</b>	<b>27</b>	<b>11</b>	<b>62</b>
<b>Web/Paper D</b>			
Previous Web Response	25	9	66
No Previous Web Response	--	--	--
<b>Total</b>	<b>25</b>	<b>9</b>	<b>66</b>
<b>Overall Web/Paper Groups</b>	<b>27</b>	<b>11</b>	<b>62</b>
<b>Main Group (control)</b>	<b>5.0</b>	<b>28.6</b>	<b>66.4</b>

These results are more in line with what was expected. The web response rates were, across the sample, much higher than is typical. We were able to pick up some additional mail responses from both the initial and follow-up mailings as well. Overall, the non-response rates were generally lower than both the Web and Main samples.

However, if these results are “typical” the implications are mixed: the primary cost saving of web returns is minimizing mail out/mail return costs. If we still have to mail out a paper questionnaire, some of our expected gains are eliminated. That said, these results do seem indicative of the magnitude of the role that having high speed internet access plays in getting online responses.

### 3.3 Main Group and Overall Rates

The final analysis of response rates was to look at the Main sample group as well as the overall rates. The Main group received a standard mail out package (stock cover letter, web reporting instructions and paper questionnaire) for both the initial and follow-up mailings.

**Figure 18:** Main sample and Overall web response, mail response and non-response rates

Treatment Group	Web Response (%)	Mail Response (%)	Non-response (%)
<b>Main Group</b>	<b>5.0</b>	<b>28.6</b>	<b>66.4</b>
<b>Web Groups</b>	<b>22</b>	<b>0</b>	<b>78</b>
<b>Web/Paper Groups</b>	<b>27</b>	<b>11</b>	<b>62</b>
<b>Overall Rates</b>	<b>13.0</b>	<b>19.6</b>	<b>67.4</b>

The Main group achieved fairly “typical” web response rates for a NASS survey. As expected, the mail response was much higher than the Web groups. This is also in line with typical voluntary NASS survey rates. The overall web and mail response rates reflect the differences encountered with the various samples and treatments.

### 3.4 Timeliness of Web Responses

The bulk of NASS survey web responses typically occur in a “bubble” of about one week, approximately one week after the mailing. This was no exception. For this test, we were primarily interested in whether a larger proportion of the web response target groups occurred in a more timely fashion.

**Figure 19:** Timeliness of web responses: Proportions occurring in web response “bubble” window

Treatment Group	Proportion in web response “bubble”
Web A	44.5%
Web B	43.5%
Web C	33.5%
Web D	48.8%
Web/Paper A	36.6%
Web/Paper B	32.4%
Web/Paper C	28.5%
Web/Paper D	40.1%
Main	24.9%

Our initial thought for the 2012 COA was to have an early mail out to high speed internet users and then include the web nonrespondents in the general mass mailing. This would require prompt response from those targeted for a web completion. From this experiment, about 25 percent of the Main sample (those not specifically targeted for web response) responses occurred in this “bubble”. The Web and Web/Paper groups ranged from just under 29 percent (in a treatment group that experienced some problems in the mailing) to almost 49 percent. These proportions are somewhat higher, but we were hoping for closer to 50 percent across all the web response target groups. The timeliness of web responses would present somewhat of a challenge in logistics and handling, as it does take some time for name and address file prep and printing prior to the mass mailing. The challenges are not insurmountable, but these results were somewhat disappointing.

### **3.5 Additional Discussion**

The primary goal of this research was to assess approaches to obtaining more web responses and the timeliness of the responses. We did not undertake much analysis of the actual reported data across the sample groups (Web, Web/Paper and Main) or the treatment groups within the two web sample groups. We also did not compare the mail responses data to the web responses data, like we did with the 2007 COA data. The overall number of responses was fairly limited and all the sample groups were selected for specific attributes: the web response groups for high speed internet access and the Main group for specific items in the questionnaire we wanted to target for testing. This would have limited the usefulness of reported data comparisons.

## **4. CONCLUSION AND RECOMMENDATIONS**

While there were some key differences between mail and web respondents in the 2007 COA, the most important factor in determining the propensity of web response was high speed internet access. In this preliminary test for the 2012 COA, we found that specifically targeting high speed internet users could indeed increase web response rates. However, for some of our treatment groups, there was a negative impact on overall response rates—respondents either reported via the web or not at all. Additionally, the timeliness of the web responses was not quite as prompt as we had initially hoped.

Approximately one-third of 2007 COA respondents reported having high speed internet access. If even 20 percent of those respond via the web through an initial web only mail out, the savings in printing, postage, key entry and data analysis would be significant. Of course, we still need to further weigh the impact on overall response rates. If this increased proportion of web response can occur without any negative repercussions on overall response rates, this approach could substantially increase our efficiency and data quality.

### **4.1 Items for Further Research**

Web reporting has many distinct benefits, including decreased costs (both in mailing and data entry), speed of data availability, and increased data quality. Given the results of this round of experiments, we are planning several more rounds of testing prior to finalizing plans for the 2012 COA. We want to further examine the impact on overall response rates when encouraging web response as well as the content of mailing packages for web response target groups. Our plans will include more emphasis on online reporting in the mailings for the census, as well as enhancements to the web instrument itself. In this way, we hope to increase the overall number of respondents reporting online.

## **5. REFERENCES**

- Biemer, Paul, and Lyberg, Lars (2003), *Introduction to Survey Quality*, New York: Wiley.
- Couper, M.P., Traugott, M.W., Lamias, M.J., (2001), Web Survey Design and Administration, *Public Opinion Quarterly*, 65, 230-253.

- Couper, M.P. (2000), "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly*, 64: 464-494.
- Dillman, D. (2000), *Mail and Internet Surveys: The Tailored Design Method*, New York: Wiley.
- Millar, M.M., O'Neill, A.C., Dillman, D.A., (2009), "Are Mode Preferences Real?" Technical Report 09-003. Washington State University Social and Economic Sciences Research Center.
- Sudman, S., Bradburn, N.M., and Schwarz, N. (1996), *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*, San Francisco: Jossey-Bass Publishers.
- Tourangeau, R., Rips, L.J., and Rasinski, K. (2000), *The Psychology of Survey Response*, Cambridge: Cambridge University Press.