UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

January 19, 2010

DSSD 2010 AMERICAN COMMUNITY SURVEY MEMORANDUM SERIES #ACS10-RE-01

MEMORANDUM FOR	Susan Schechter Chief, American Community Survey Office
From:	David C. Whitford ***/signed/*** Chief, Decennial Statistical Studies Division
Prepared by:	Megha Joshipura ***/signed/*** American Community Survey Data Collection Methods Branch
Subject:	Analysis of the 2009 Field of Degree Question in the American Community Survey

Attached is the final report for the Analysis of the 2009 Field of Degree Question in the American Community Survey. This report summarizes the results from evaluating the new field of degree question.

If you have any questions about this report, please contact Megha Joshipura (x31643).

Attachment: (Analysis of the 2009 Field of Degree Question in the American Community Survey)

cc: Todd Hughes (ACSO)
David Raglin
Anthony Tersine (DSSD)
Jennifer Guarino Tancreto
Mary Frances Zelenak
Padraic Murphy
Jennifer Karberg
Robert Kominski (HHES)
Kurt Bauman

Intentionally Blank

Analysis of the 2009 Field of Degree Question in the American Community Survey

EVALUATION REPORT

Megha Joshipura
Decennial Statistical Studies Division

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

TABLE OF CONTENTS

1.0 Background

The field of degree question is new to the 2009 American Community Survey (ACS) and the 2009 Puerto Rico Community Survey (PRCS). Two versions of the question were tested in the 2007 ACS Content Test, an open-ended version and a categorical version. The open-ended version performed better according to the evaluation criteria set forth for the content test (Raglin *et. al.*, 2008). The purpose of this evaluation is to assess how the field of degree question is performing in production on both the ACS and the PRCS.

On the 2009 ACS and PRCS English paper questionnaires, the field of degree question is located directly below the educational attainment question in the same column. However, on the paper questionnaires created for the 2007 ACS Content Test, the educational attainment question is at the bottom of one column while the field of degree question is at the top of the next column. The question was tested this way because the categorical version of the question was longer and would not fit under the educational attainment question. Since the ACS wanted to compare both questions in the same placement on the test, both questions were tested at the top of the next column. We believe that this difference could cause differences in item nonresponse between the English versions of the questionnaires in the 2007 ACS Content Test and the 2009 ACS production. On the 2009 ACS and PRCS Spanish paper questionnaires, the educational attainment question is also at the bottom of one column while the field of degree question is at the top of the next column because of space constraints.

2.0 Research questions

In order to evaluate the field of degree question, we looked at the ACS Housing Units (HU), the ACS Group Quarters (GQ), and the PRCS individually. The PRCS only includes HUs because there were no GQs residents with a Bachelor's degree in sample.

For the ACS HU, we looked at the following research questions:

- Is the distribution of responses to the field of degree question in the 2009 ACS HU comparable to the open-ended version of the question that was tested in the 2007 ACS Content Test?
- Are the relative distributions of field of degree categories roughly comparable to existing data sources, such as the National Survey of College Graduates (NSCG)?
- Is the item nonresponse rate to the field of degree question in the 2009 ACS HU comparable to the open-ended version of the question that was tested in the 2007 ACS Content Test?
- Are the responses in the 2009 ACS HU as consistent with the educational attainment question as the responses in the 2007 ACS Content Test?
- Is there a difference in the percent auto-coded versus clerically-coded between the 2009 ACS HU and the 2007 ACS Content Test?

For the ACS GQ, we looked at the following research questions:

- What is the item nonresponse rate for the field of degree question in the 2009 ACS GQ?
- How consistent are the responses with the educational attainment question on the questionnaire?

For the PRCS, we looked at the following research questions:

- What is the item nonresponse rate for the field of degree question in the 2009 PRCS?
- How consistent are the responses with the educational attainment question on the questionnaire?
- What is the percent auto-coded versus clerically-coded? Is it comparable to the 2009 ACS?

3.0 Methodology

This evaluation examines the field of degree question in the ACS HU, ACS GQ, and the PRCS. The ACS and PRCS HU collect data in three modes over three months (U.S. Census Bureau, 2009). During the first month, mail questionnaires are sent to each of the mailable addresses in sample for that month. During the second month, the previous months' non-responding addresses that have a listed telephone number are called in an attempt to get a computer-assisted telephone interview (CATI). During the third month, we take a sub-sample of all mailable addresses that have not responded in mail or CATI and a sub-sample of all unmailable addresses and try to get a computer-assisted personal interview (CAPI). Questionnaires that are sent back in the mail are checked-in and sent for keying after we receive them. After they are keyed, the questions that need to be coded, such as the field of degree question, are sent for coding. The CATI and CAPI interviews are keyed as the interview is being taken so they are sent directly for coding. The ACS and PRCS GQ are primarily conducted through interviews by field representatives (FR), either by telephone or personal interview. Most of the interviews are actually taken on an automated instrument. While it is usually completed with an FR present, there are some cases where paper questionnaires are left for the sample residents to fill out on their own. After the interviews are conducted, the questionnaires are sent back to be checked-in, keyed, and coded, if necessary.

For the ACS HU, this evaluation entails comparing the production field of degree data with the results of the 2007 ACS Content Test. We included both English and Spanish forms and did not distinguish between the two because of the small number of Spanish forms sent out for the ACS and small number of English forms sent out for the PRCS. We excluded Alaska and Hawaii from our comparisons to the 2007 ACS Content Test because the test did not include them. Because the field of degree question was only tested in the ACS HU, we could not compare

the ACS GQ or the PRCS to the results of the 2007 ACS Content Test. We looked at field of degree responses in all modes for data collected in January 2009 (excluding mail returns from the November and December 2008 panels) through May 2009. For GQs, this includes panels assigned for Remote Alaska. All analyses were conducted at the national level. We also conducted a subset of the analyses at the state level. For the ACS HU, we did not do any editing or imputation on the data so as not to mask any response problems, and to maintain comparability to 2007 ACS Content Test data that were also not edited or imputed. We also did not edit or impute any data in PRCS and GQ in order to see how respondents truly answer the field of degree question.

This evaluation looks at the distribution of responses by category, item nonresponse rates, and whether or not the question was answered consistently. To calculate the distribution of response by category, the following formula was used:

$$\text{Response Distribution} = \frac{\#\text{ of codable responses in category}}{\#\text{ of respodents that gave a codable response}}$$

We calculated the item nonresponse rates in two ways. First, we only considered cases where the respondent did not respond to the field of degree (FOD) question as item nonresponses:

$$\text{Item Nonresponse Rate 1} = \frac{\#\text{ of respondents that did not respond to the FOD question}}{\#\text{ of respondents with a Bachelor's degree}}$$

Then, we considered those cases and the cases where we could not code the responses as item nonresponses:

$$\text{Item Nonresponse Rate 2} = \frac{\#\text{ of respondents that did not respond to the FOD question or gave an uncodable answer}}{\#\text{ of respondents with a Bachelor's degree}}$$

Whether or not the question was answered consistently was calculated using the formula below. By consistency, we mean if the respondent answered the question when they were supposed to (as in, they have a Bachelor's degree) or did not answer the question when they were not supposed to.

$$\text{Inconsistency rate} = \frac{\#\text{ of respondents that answered the FOD question inconsistently}}{\#\text{ of respondents who answered educational attainment}}$$

4.0 Limitations

We did not edit the data that respondents reported, nor did we impute for missing responses. We also did not adjust the weights for unit nonresponse or to match known population totals, as is done with ACS production data. We purposely chose not to implement these adjustments in order to prevent masking any response problems the question might have. This is a limitation in that the data cannot be easily compared to published data sources, such as the NSCG. However, we did so to make the data comparable to the content test data. There is also a different coding system in place for the 2009 ACS than there was for the 2007 Content Test.

5.0 Results

ACS Housing Units

5.1 Is the distribution of responses to the field of degree question in the 2009 ACS HU comparable to the open-ended version of the question that was tested in the 2007 ACS Content Test?

Yes, the distribution of responses to the field of degree question in the 2009 ACS HU is comparable to the open-ended version of the question that was tested in the 2007 ACS Content Test. There are some significant differences. However, this could be because of the difference in the coding systems. We should note that, in 2009, out of all of the responses provided by the respondents, approximately 5% of the responses were uncodable[1]. In 2007, approximately 3.5% of the responses were uncodable.

Table 1 shows the distribution of the percentage of people in each field of degree category. The totals for both the 2009 ACS HU and the 2007 ACS Content Test are both over 100 percent, meaning that some people were reported to have degrees in more than one category. The distributions below do not include the uncodable answers. In 2009, the weighted number of respondents that gave a codable answer was 20,761,388. In 2007, that number was 46,383,500.

[1] An uncodable response means that the response could not be categorized into the list of bachelor's degree fields. For example, 'vet medicine' is considered to be uncodable because it refers to a type of degree that is not offered at the bachelor's level.

Table 1. Comparison of Percent Distributions of Field of Degree Category (standard errors in parentheses)

Field of Degree Category	2009 ACS HU	2007 ACS Content Test	Difference
Biological, agricultural, physical, earth, or other natural sciences	8.3 (0.04)	9.9 (0.55)	-1.6* (0.55)
Health, nursing, or medical fields	7.2 (0.05)	7.6 (0.47)	-0.4 (0.47)
Engineering, computer sciences, or mathematical sciences	13.7 (0.08)	12.5 (0.61)	1.1* (0.62)
History, arts, or humanities	15.9 (0.07)	14.8 (0.66)	1.1 (0.67)
Psychology, economics, or other social sciences	14.5 (0.07)	15.2 (0.59)	-0.7 (0.60)
Business or management	21.6 (0.08)	22.4 (0.86)	-0.8 (0.86)
Education or education administration	15.1 (0.06)	14.0 (0.63)	1.1* (0.64)
Some other major field	10.9 (0.07)	10.1 (0.62)	0.8 (0.62)

*Significant at the .10 alpha level.

5.2 Are the relative distributions of field of degree categories roughly comparable to existing data sources, such as the NSCG?

The 2003 NSCG is the best source of data on a bachelor's field of degree category. The 2003 NSCG estimated the number of people with a bachelor's degree to be 40,036,000. While we cannot do any formal statistical comparisons between the 2009 ACS field of degree question and the 2003 NSCG because of differences in data processing, the relative distributions of the 2009 ACS HU seem roughly comparable to the 2003 NSCG.

Table 2. Comparison of Percent Distributions of Field of Degree Category to the NSCG

Field of Degree Category	2009 ACS HU	2003 NSCG
Biological, agricultural, physical, earth, or natural sciences	8.3	9.3
Health, nursing, or medical fields	7.2	7.2
Engineering, computer sciences, or mathematical sciences	13.7	13.2
History, arts, or humanities	15.9	15.5
Psychology, economics, or other social sciences	14.5	14.2
Business or management	21.6	20.9
Education or education administration	15.1	15.2
Some other major field	10.9	8.2

5.3 Is the item nonresponse rate to the field of degree question in the 2009 ACS HU comparable to the open-ended version of the question that was tested in the 2007 ACS Content Test?

For this research question, we only looked at respondents that were eligible to respond to the field of degree question. That is, respondents who had a Bachelor's degree. We calculated the item nonresponse rates in two ways. First, we only considered cases where the respondent did not respond to the field of degree question as item nonresponses. Then, we considered those cases and the cases where we could not code the responses as item nonresponses. In 2009, the weighted number of respondents with a Bachelor's degree was 22,571,308. In 2007, the number was 49,598,381. Both item nonresponse rates to the field of degree question in the 2009 ACS HU are significantly lower than those in the 2007 ACS Content Test. This could possibly be due to the different placements on both forms. The 2009 ACS HU includes telephone questionnaire assistance (TQA) and failed-edit follow up (FEFU)[2], which could also be another difference since the 2007 ACS Content Test did not include these.

Table 3. Comparison of Item Nonresponse Rates (standard errors in parentheses)

	2009 ACS HU	**2007 ACS Content Test**	**Difference**
Nonresponse Only	1.3	3.5	-2.1*
	(0.02)	(0.33)	(0.33)
Uncodable+Nonresponse	5.8	6.5	-0.7
	(0.05)	(0.45)	(0.45)

*Significant at the .10 alpha level.

There is also a significant increase in the item nonresponse rate when uncodable responses are taken into account and the correlation between the two item nonresponse rates is accounted for.

Table 4. Two Types of Item Nonresponse Rates on the 2009 ACS HU (standard errors in parentheses)

	Nonresponse Only	**Uncodable + Nonresponse**	**Difference**
Item nonresponse rate	1.3	5.8	-4.5*
	(0.02)	(0.05)	(0.04)

*Significant at the .10 alpha level.

We also looked at item nonresponse by mode. The mail item nonresponse rate is significantly higher than both CATI and CAPI. CAPI also has a slightly higher item nonresponse rate than CATI does. When uncodable responses are included, there are no significant differences between mail and CATI. However, the CAPI item nonresponse rate significantly increases and becomes significantly higher

[2] TQA provides assistance to respondents over the telephone. It also allows the respondent to complete the survey over the telephone, if they prefer. The FEFU operation calls respondents who did not complete all of the critical items on the survey and finishes the survey over the phone. Both of these help decrease incidents of unit and item nonresponse.

than mail or CATI. This result is interesting, in that it means that there are a higher percentage of uncodable cases for CATI and CAPI than there is for mail. In fact, both CATI and CAPI have a significantly higher number of uncodable responses than mail does. CAPI also has a significantly higher number than CATI. This suggests that the reason for the high number of uncodable responses might have to do with some interviewer keying error.

Table 5. Item Nonresponse Rates by Mode for the 2009 ACS HU (standard errors in parentheses)

Mode	**Nonresponse only**	**Uncodable + Nonresponse**	**Uncodable only**
Mail	1.9	5.1	3.3
	(0.03)	(0.05)	(0.04)
CATI	0.0	5.1	5.1
	(0.01)	(0.14)	(0.14)
CAPI	0.1	8.3	8.2
	(0.02)	(0.15)	(0.14)
Difference between	1.8*	0.0	-1.8*
Mail and CATI	(0.03)	(0.15)	(0.14)
Difference between	1.8*	-3.2*	-5.0*
Mail and CAPI	(0.03)	(0.16)	(0.15)
Difference between	-0.1*	-3.3*	-3.2*
CATI and CAPI	(0.02)	(0.20)	(0.20)

*Significant at the .10 alpha level. Multiple comparisons were accounted for using the Bonferroni method.

We also looked at item nonresponse by state and compared each state to the overall item nonresponse rate. The states that have an item nonresponse rate that is significantly different from the overall item nonresponse rate are marked with an asterisk in the table below. Each test was done at a 10 percent significance level, with correlation taken into account. No adjustments were made for multiple comparisons.

Table 6. Item Nonresponse Rates by State for the 2009 ACS HU (standard errors in parentheses)

State	Nonresponse only	Uncodable + Nonresponse
Alabama	1.3	5.9
	(0.18)	(0.45)
Alaska	0.5*	5.9
	(0.20)	(1.04)
Arizona	1.1	5.7
	(0.14)	(0.38)
Arkansas	0.9*	4.3*
	(0.18)	(0.53)
California	1.5*	6.8*
	(0.06)	(0.14)
Colorado	0.9*	4.4*
	(0.12)	(0.29)
Connecticut	1.1*	5.7
	(0.16)	(0.41)
Delaware	0.7*	4.1*
	(0.24)	(0.68)
District of Columbia	1.3	6.7
	(0.31)	(1.01)
Florida	1.8*	7.8*
	(0.10)	(0.24)
Georgia	1.3	5.2*
	(0.13)	(0.27)
Hawaii	1.0	6.0
	(0.24)	(0.76)
Idaho	0.9*	4.2*
	(0.25)	(0.64)
Illinois	1.4	6.0
	(0.09)	(0.23)
Indiana	1.1*	4.0*
	(0.12)	(0.29)
Iowa	1.0	3.5*
	(0.18)	(0.36)
Kansas	0.9*	4.1*
	(0.15)	(0.34)
Kentucky	1.2	5.0*
	(0.17)	(0.42)
Louisiana	1.4	5.7
	(0.16)	(0.34)
Maine	1.2	4.3*
	(0.28)	(0.51)
Maryland	1.3	5.6
	(0.11)	(0.33)
Massachusetts	1.6*	6.0
	(0.13)	(0.32)
Michigan	1.5	5.8
	(0.11)	(0.30)
Minnesota	1.4	4.5*
	(0.14)	(0.28)
Mississippi	1.3	5.4
	(0.20)	(0.60)
Missouri	1.2	4.8*
	(0.13)	(0.33)

State	Nonresponse only	Uncodable + Nonresponse
Montana	1.4	3.4*
	(0.31)	(0.56)
Nebraska	1.4	4.4*
	(0.28)	(0.48)
Nevada	1.3	7.1*
	(0.20)	(0.69)
New Hampshire	0.9*	4.4*
	(0.19)	(0.56)
New Jersey	1.5	6.8*
	(0.10)	(0.30)
New Mexico	0.9*	4.9
	(0.22)	(0.61)
New York	1.5	7.9*
	(0.09)	(0.21)
North Carolina	1.2	4.6*
	(0.10)	(0.25)
North Dakota	1.3	4.6
	(0.43)	(0.84)
Ohio	1.3	5.4*
	(0.09)	(0.27)
Oklahoma	1.3	5.3
	(0.19)	(0.48)
Oregon	1.4	5.1*
	(0.18)	(0.40)
Pennsylvania	1.2*	4.8*
	(0.09)	(0.22)
Rhode Island	1.4	6.4
	(0.30)	(0.88)
South Carolina	1.3	5.1*
	(0.18)	(0.42)
South Dakota	1.6	6.0
	(0.30)	(0.81)
Tennessee	1.4	5.3
	(0.15)	(0.41)
Texas	1.2	5.9
	(0.07)	(0.21)
Utah	0.7*	4.3*
	(0.13)	(0.51)
Vermont	1.2	5.2
	(0.33)	(0.79)
Virginia	1.2	5.4*
	(0.09)	(0.25)
Washington	1.0*	4.6*
	(0.12)	(0.35)
West Virginia	1.5	6.9
	(0.27)	(0.85)
Wisconsin	1.2	4.3*
	(0.15)	(0.26)
Wyoming	1.5	3.7*
	(0.49)	(0.83)

*Significantly different from the overall rates at the .10 alpha level.

5.4 Are the responses in the 2009 ACS HU as consistent with the educational attainment question as the responses in the 2007 ACS Content Test?

Here, consistency means whether the respondent correctly answered or skipped the question based on their answer to the educational attainment question. To answer this question, we only looked at the mail respondents because the programmed skips in the CATI/CAPI instrument do not allow for inconsistency. The field of degree responses are more consistent with the educational attainment question in the 2009 ACS HU than in the 2007 ACS Content Test. This may be due to the different placements of the field of degree question on the form. Because there is no break between the educational attainment question and the field of degree question, respondents seem to be more likely to answer the question correctly.

Table 7. Comparison of the Percent of Inconsistent Answers (standard errors in parentheses)

	2009 ACS HU	**2007 ACS Content Test**	**Difference**
Percent of Inconsistent Answers	0.8 (0.01)	1.9 (0.14)	-1.1* (0.14)

*Significant at the .10 alpha level.

5.5 Is there a difference in the percent auto-coded versus clerically-coded between the 2009 ACS HU and the 2007 ACS Content Test?

There is a difference in the percent auto-coded versus clerically-coded between the 2009 ACS HU and the 2007 Content Test. The 2009 ACS HU has a much higher percent auto-coded because we used a different coding system in 2009 that included responses collected in the 2007 Content Test. Also, the 2009 coding system puts the clerically-coded responses from the earlier months of 2009 into the system for later months. In January and February, the percent auto-coded was 66 percent, which was already higher than the 57 percent auto-coded in 2007. By May 2009, the percent auto-coded increased to 76 percent.

ACS Group Quarters

5.6 What is the item nonresponse rate for the field of degree question in the 2009 ACS GQ?

For this research question, we only looked at respondents that were eligible to respond to the field of degree question (those with a Bachelor's degree). For the 2009 ACS GQ, the weighted number of GQ respsondents with a Bachelor's degree was 21,118. The item nonresponse rate was 12.1 percent, when looking at nonresponse only. When uncodable responses were included, there was a small but significant increase in the item nonresponse rate. While the item nonresponse rates might appear to be high here, they seem to be in the range of other item nonresponse rates in the ACS GQ. The educational attainment question, for example, has an item nonresponse rate of 9.8 percent. A potential reason for

these high item nonresponse rates could be the use of proxy interviews in the ACS GQ.

Table 8. Comparison of Item Nonresponse Rates on the 2009 ACS GQ (standard errors in parentheses)

	Nonresponse only	Uncodable + Nonresponse	Difference
Item nonresponse rate	12.1	14.1	-2.0*
	(1.48)	(1.56)	(0.51)

*Significant at the .10 alpha level.

5.7 How consistent are the responses with the educational attainment question on the questionnaire?

Only 1.6 percent of all respondents to the ACS GQ answered this question inconsistently. Therefore, it seems respondents are answering this question consistently with the educational attainment question. One of the reasons for this percent to be so small is that the preferred method of response for ACS GQ is for the interviewer to administer the survey via an automated instrument.

Puerto Rico Community Survey

5.8 What is the item nonresponse rate for the field of degree question in the 2009 PRCS?

For this research question, we only looked at respondents that were eligible to respond to the field of degree question (those with a Bachelor's degree). For the 2009 PRCS, the weighted number of respondents with a Bachelor's degree was 185,330. The item nonresponse rate was 0.6 percent, when looking at nonresponse only. When uncodable responses were included, there was a significant increase in the item nonresponse rate. The high number of uncodable responses shows that there might be a need to examine the translation of the question and to make sure respondents understand it correctly.

Table 9. Comparison of Item Nonresponse Rates for the 2009 PRCS (standard errors in parentheses)

	Nonresponse only	Uncodable + Nonresponse	Difference
Item nonresponse rate	0.6	9.5	-8.9*
	(0.13)	(0.66)	(0.64)

*Significant at the .10 alpha level.

5.9 How consistent are the responses with the educational attainment question on the questionnaire?

Only 1.6 percent of all mail respondents to the PRCS answered this question inconsistently. Therefore, it seems respondents are answering this question consistently with the educational attainment question.

5.10 What is the percent auto-coded versus clerically-coded? Is it comparable to the 2009 ACS?

The percent auto-coded for the PRCS is approximately 60 percent. The major difference between the ACS and the PRCS is that the percent auto-coded has increased greatly as the year goes on for the ACS, but has stayed stable around 60 percent for the PRCS. The reason for this could be that PRCS uses a bilingual coding operation, which was not used in the 2007 ACS Content Test, so their operation is newer than the one used on the state-side forms.

6.0 Conclusions

The new field of degree question seems to be working as expected in the 2009 ACS and PRCS. The distribution of responses is similar to that found in the 2007 ACS Content Test. The 2009 ACS HU outperformed the 2007 ACS Content Test in terms of item nonresponse rate and consistency, which is most likely because of the new placement of the question. Also, the new coding system is working very well for the ACS and the PRCS. However, the high percentage of uncodable answers coming from the CATI and CAPI modes of data collection should be investigated in more detail.

7.0 References

Raglin, D., Zelenak, M., Davis, M., and Tancreto, J. (2008) "Testing a New Field of Degree Question". 2007 American Community Survey Content Test Report, May 2008.

U. S. Census Bureau (2009), *Design and Methodology*, American Community Survey. U.S. Government Printing Office, Washington, DC, 2009.