Using Names to Improve Measurement of Same-sex Married Couples

in the American Community Survey

Author: Daphne Lofquist

Affiliation: U.S. Census Bureau, Social, Economic, and Housing Statistics Division

Conference: Annual meeting of the Population Association of America,

San Diego, CA, April 30 - May 2, 2015

Form: Poster extended abstract

Abstract:

The U.S. Census Bureau conducts ongoing research to improve measurement of all household relationships, especially for same-sex married couples. Since they are a relatively small population among all households, same-sex married couple households are difficult to estimate accurately. Issues discovered with the 2010 Census led to the development of "preferred estimates" of same-sex couples. Using the same-sex married couple flag and the names index, this poster extends prior research by using data from the 2013 American Community Survey to identify those couples who likely are same-sex married couples compared to those who are most likely opposite-sex couples who mismarked the sex item for at least one of the spouses. These results will help further the Census Bureau's research on same-sex married couples and achieve our goal of producing high-quality estimates.

Marital and household relationships in the United States have become more complex since the latter half of the 20th century. This is especially true for same-sex couples. Data from the U.S. Census Bureau are used in a variety of applications including research on family change, stability, and instability. Although American family life has changed, measures of household and family relationships have failed to keep pace. Demographic surveys must adapt measures that reflect the complexity of contemporary relationships so that they can accurately portray American households and families. The Census Bureau continues to work to improve measurement of same-sex unmarried and married couples.

Currently, 37 states (or state equivalents) recognize same-sex marriage.[1] The June 2013 Supreme Court ruling on the Defense of Marriage Act (DOMA) heightened the need for research on this topic. Since same-sex marital and nonmarital relationships are receiving greater recognition, researchers have started to focus on how same-sex couples report their relationship and marital status on demographic surveys.

Using the 2013 American Community Survey, the current study uses the same-sex married couple flag and the names index to identify those couples who likely are same-sex married couples compared to those who are most likely opposite-sex couples who mismarked the sex item for at least one of the spouses.

BACKGROUND

The Census Bureau conducts ongoing research to improve measurement of all household relationships, especially for same-sex married couples. Since they are a relatively small population among all households, same-sex married couple households are difficult to estimate accurately. A small rate of

[1] These are as follows: Alaska, Arizona, California, Colorado, Connecticut, Delaware, Florida, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Maine, Maryland, Massachusetts, Minnesota, Missouri, Montana, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin, and Wyoming and the District of Columbia.

error in the large group (opposite-sex married couples) can result in a large amount of error in the small group (same-sex married couples). In the 2010 Census, this issue was found to affect the measurement of same-sex couple households. After discovering that the estimates of same-sex couple households in the 2010 Census data were likely inflated by mismarks on the sex question by opposite-sex couples, Census Bureau researchers used a first names index to adjust the estimates. The adjusted estimates are referred to as the "preferred estimates" in published materials, and the methodology is described in detail in the O'Connell and Feliz (2011) report.

Subsequent research included the testing of additional answer categories for the relationship question, which explicitly ask if the respondent is the "opposite-sex husband/wife/spouse" or "same-sex husband/wife/spouse" of the householder. Additional research matched Social Security data with American Community Survey (ACS) records and 2010 Census data, which confirmed that there continues to be a problem with opposite-sex married couples misreporting sex and subsequent inflation of the estimates of same-sex married couples (Kreider and Lofquist 2015). The problem was not as pronounced in the ACS as it was in the 2010 Census, but was still sizable. Kreider and Lofquist found that assigning sex based on the first name, as was done for the "preferred estimates" released from 2010 Census data, was a fairly effective way to adjust the estimates to get a more accurate sense of those who reported being same-sex married couple households in the decennial census. The Census Bureau has found that there is still a high proportion of same-sex married couples, whose reports of sex indicate they are likely opposite-sex married couples, who have marked sex or the same-sex relationship category in error (Lofquist and Lewis forthcoming).

DATA

American Community Survey

The analyses in this poster use the relationship and sex items in the 2013 ACS. The ACS was fully implemented in 2005 and was designed to replace the collection of data from the long-form decennial census questionnaire that was previously distributed to 1 in 6 households in Census 2000. The ACS is a mandatory survey that is conducted annually over a 12-month calendar period. The Census Bureau mails approximately 250,000 ACS questionnaires every month to a nationwide sample. The questionnaire is administered through a mailout/mailback paper form and an internet mode (introduced in 2013), with a nonresponse follow up using computer-assisted telephone and/or in-person interviews. The final unweighted sample is approximately 2.2 million U.S. households in 2013.[2] This sample is then weighted to be representative of the nation's population as a whole. The ACS provides nationally representative data on households, which includes social, demographic, economic, and housing data. Given its large sample size, the ACS is one of the only surveys large enough to reliably estimate small populations like same-sex married couples.[3]

The ACS collects information on the relationship of each member of the household to the householder (the person who owns or rents the home). The category "unmarried partner" has been on the American Community Survey since its full implementation in 2005. Same-sex unmarried partners were first reported in the 1990 Census, and the 2010 Census marked the first published estimates based on decennial data of those who identified themselves as same-sex married couples. Data from Census 2000 reported all same-sex couples as unmarried couples, as no states performed same-sex marriages at that

[2] http://www.census.gov/acs/www/methodology/sample_size_data/index.php
[3] Additional information about the ACS, its methodology, and data products can be found at <http://www.census.gov/acs/www>.

time. The Census Bureau has also released yearly estimates of those who reported as same-sex married couple households in ACS going back to 2005.[4]

Data are edited to correct inconsistencies and protect respondent confidentiality. The procedure for editing the responses of those who report being married to a same-sex partner has changed over time. In public use ACS data for 2005 to 2012, in cases where no imputations were made due to non-response on either the person's relationship or gender, a same-sex partner who reported being a "spouse" of the householder was changed to an "unmarried partner" of the householder.

Starting with the 2013 ACS 1-year data file, same-sex spouses are no longer edited to be same-sex unmarried partners. This change to the edit not only includes those same-sex spousal households in which, for either person, the relationship and gender items were not missing, but it also includes those couples where either the householder or spouse did not report their sex, and it was assigned based on their first name.[5] A flag is provided on the Public Use Microdata Sample (PUMS) data set that identifies these two groups of same-sex married couples. These couples are shown in tables in the categories of labeled 'husband or wife' or 'married couple households.' Same-sex and opposite-sex spouses are grouped together in tabulations rather than shown separately, due largely to the small relative size of the same-sex group.

The names index, like the one used by O'Connell and Feliz (2011), is available on the ACS internal data set. The index is based on 2010 Census reports of names by state. This index created tallies of the number of times a particular name was reported as male. This index value can be used to assign the probable sex of the record holder. If a name was reported as male 95-percent of the time, then the report of sex for that person was set to male. The same was done for names reported as female. An

[4] These estimates are available in a set of tables titled "Characteristics of Same-Sex Couple Households," available at: http://www.census.gov/hhes/samesex/data/acs.html

[5] Missing data includes those with relationship and/or sex missing or were had their data allocated for some other reason, which could include those whose response was "Don't Know" or "Refused."

"ambiguous" name (labeled "cannot determine" in the accompanying poster) means that the sex could not be assigned based on the 95-percent rule. This included names that were male less than 95 percent, but more than 5 percent of the time, as well as names that were unique, or cases in which the name field was blank. In addition, if a name was not reported at least 10 times, then the index value cannot be determined.

Using the same-sex married couple flag and the names index, this poster extends prior research by O'Connell and Feliz (2011) and by Lofquist and Lewis (forthcoming) by using the data from the 2013 American Community Survey to identify those couples who likely are same-sex married couples compared to those who are most likely opposite-sex couples who mismarked the sex item for at least one of the spouses. The adjustment to the edit in 2013 means some same-sex married couples in the edited data have not reported all of the information that defines them as such. Because of this, I want to investigate the characteristics of those who have missing data, to see whether they are likely to be same-sex married couples, or whether the edit may be assigning them as such in error.

Research Goals

1. Evaluate the change to the relationship edit for those who have missing data by comparing those with missing data to those without missing data.
2. Using the names index, identify those couples who are likely same-sex married couples.
3. Compare those who are likely same-sex married couples to those who are most likely not same-sex married couples.

RESULTS

Table 1 shows the total weighted number of same-sex married couples by presence of missing relationship and/or sex data. This table shows that there were 251,695 same-sex married couples in the 2013 ACS. Of these couples, 15,440, or 6 percent, had at least one spouse who had a missing value on the relationship and/or sex item. Approximately 17 percent of mailout/mailback respondents were

missing relationship and/or sex data compared to the 1 percent of CATI/CAPI respondents and less than half a percent of those who responded using the internet.

Table 2 shows characteristics of same-sex married couples by whether or not relationship and/or sex were missing. Householders and their spouses with missing data tended to be older than those with no missing data. A smaller percentage of couples with missing data had a householder with at least a Bachelor's degree than those with no missing data. This is especially pronounced for couples where both spouses have at least a Bachelor's degree. Those couples with no missing data are more likely to have a White householder or to be in an interracial relationship than those with missing data. Those with no missing data have a much higher average household income than those with missing data ($116,500 and $83,505, respectively).

Table 3 shows the distribution of same-sex married couples by whether or not their edited sex matches the sex in the names index. Couples are considered to be "likely same-sex married couples" when both the householder's sex and spouse's sex matches the sex they would be assigned using the names index. Those who do not match are listed as "not likely same-sex married couple," and if one or both spouses' sex cannot be determined using the names index they are shown under "Cannot determine. " Approximately 68 percent of same-sex married couples in our data are likely same-sex married couples. This compares to about 16 percent who are not likely a same-sex married couple and approximately 17 percent for whom we cannot determine whether they are a same-sex married couple. For those with missing data, only 56 percent were likely same-sex married couples while approximately 35 percent of them could not have their sex assigned using the 95-percent rule. For those who responded using the internet, only 8 percent were likely not same-sex married couples compared to 26 percent of those who responded using the mailout/mailback mode. The corresponding percentage for those who responded using CATI/CAPI is approximately 17 percent.

Table 4 shows characteristics of same-sex married couples by whether relationship and sex data are missing and whether their name matches their reported sex.[6] Householders and spouses who are likely same-sex married couples are younger than those who are not same-sex married couples. A higher percentage of likely same-sex married couples have a householder and spouse with at least a Bachelor's degree than are those who likely are not same-sex married couples. Those couples who are likely same-sex married couples have an average household income of $123,299 while those who are not likely same-sex married couples only have an average household income of $86,007. Likely same-sex married couples are more likely to be White than those who are not same-sex married couples. They are also less likely to have a householder of Hispanic or Latino origin than those who are not likely same-sex married couples.

Same-sex couples who are missing data, for both those likely and not likely to be same-sex married couples, tend to be older than those with no missing data. They also have a smaller percentage where the householder and spouse both have at least a Bachelor's degree. These groups also report lower average household incomes. A lower percentage of those with missing data, but who are likely same-sex married couples, were White and a higher percentage were Black compared to those who are likely same-sex married couples with no missing data.

CONCLUSIONS

Measuring relationships among household members is complex, especially when measuring same-sex married couples. There were 251,695 same-sex married couples in the 2013 ACS, with relationship and/or sex allocated in 15,440 (or 6 percent) of cases. A higher percentage of same-sex married couples

[6] Standard errors for Table 4 can be found in Table 4a.

who filled out the paper form were missing the relationship and/or sex of the householder and/or spouse— approximately 17 percent, than CATI/CAPI and internet respondents.

Overall, the results show that of the 251,695 same-sex married couples in the ACS 2013, approximately 68 percent of them are likely same-sex married couples. Since the edit assigns sex based on first name for those with missing data, the apparent error rate is lower for those with missing data. However, 35 percent do not have a name that was reported as male or female 95 percent of the time, or did not provide a name. Thus, a larger proportion of those with missing data fall into the group labeled "cannot determine," which means we are less certain that they have been assigned an appropriate sex value.

Of the 251,695 same-sex married couples, 170,429 couples are likely same-sex married couples. The smallest level of apparent error is for those who responded using the internet compared to both CATI/CAPI and mailout/mailback. This result is encouraging since we are moving toward using the internet mode for survey response. However, this may simply reflect a different group of respondents who use the internet. The CATI/CAPI mode had a built in check to let the field representative (FR) know when a respondent has chosen "Husband or wife" and the spouse's sex matches the householder's sex. The check asks the FR to verify that these data are correct. This check is not present in the internet mode. This is interesting as there is a lower percentage of same-sex married couples who have missing data in the internet mode than in the CATI/CAPI. Cognitive testing has found that respondents reported sometimes thinking of the wrong person when answering questions, so seeing the exact question on the internet mode screen may help respondents to answer for the correct person more easily than hearing the name read to them during the CATI/CAPI mode.

The results also allow us to identify differences between who we think are most likely same-sex married couples versus those which are likely opposite-sex married couples. Those who responded using the

mailout/mailback form are more likely to have missing data. Those with missing data tend to be older, not White, and do not have a mortgage.

These results will help further the Census Bureau's research on same-sex married couples and continue toward our goal of producing high-quality estimates of same-sex married couples. This poster investigates mismarks on sex in the current ACS, which uses a relationship question which has the answer category "husband/wife/spouse." The Census Bureau is currently testing a revised relationship question with explicit answer categories for opposite-sex and same-sex spouses. The pattern of error may differ when data collection employs this new question wording. We will continue to evaluate how well the relationship question functions, as well as the effects of editing procedures.

REFERENCES

Kreider, Rose M. and Daphne A. Lofquist. 2015. “Matching Survey Data with Administrative Records to Evaluate Reports of Same-sex Married Couple Households.” SEHSD Working Paper #2014-36, U.S. Census Bureau: Washington, DC.

Lofquist, Daphne A. and Jamie M. Lewis, forthcoming. ”Improving Measurement of Same-sex Couples.” SEHSD Working Paper. U.S. Census Bureau: Washington, DC.

O’Connell, Martin and Sarah Feliz. 2011. “Same-Sex Couple Household Statistics from the 2010 Census.” *SEHSD Working Paper*, 2011-26. U.S. Census Bureau: Washington, DC.

Table 1. Same-sex Married Couples by Presence of Missing Data by Mode

Mode	Total		Missing relationship and/or sex[1]		No missing data	
	Estimate	Std Error	Percent	Std Error	Percent	Std Error
Total (number)	251,695	4,169	15,440	998	236,255	3,996
Total (percent)	100.0	0.0	6.1	0.4	93.9	0.4
Mailout/Mailback	87,867	2,282	16.7	1.0	83.3	1.0
CATI/CAPI	43,690	2,776	1.2	0.5	98.8	0.5
Internet	120,138	2,056	0.2	0.1	99.8	0.1

[1] Missing data could includes those with relationship and/or sex missing or had their data allocated for some reason, which includes "Don't Know" or "Refused."

Source: U.S. Census Bureau, 2013 American Community Survey 1-year data file.

Table 2. Characteristics of Same-sex Married Couples by Presence of Missing Data

Characteristics	Total		Missing relationship and/or sex[1]		No missing data	
	Percent	Std Error	Percent	Std Error	Percent	Std Error
Total (number)	251,695	4,169	15,440	998	236,255	3,996
Average age of householder (years)	53.9	0.3	60.2	1.1	53.5	0.3
Average age of spouse (years)	52.7	0.3	59.3	1.0	52.3	0.3
Average household size (number)	2.8	--	2.6	0.1	2.8	0.0
Householder with at least a BA	45.2	0.9	22.8	2.8	46.7	0.9
Both spouses with at least a BA	30.2	0.8	10.4	1.9	31.5	0.8
Lives in the midwest	17.9	0.6	25.3	2.7	17.4	0.6
Race of householder						
White alone	83.0	0.7	75.3	3.0	83.5	0.8
Black alone	7.5	0.5	13.8	2.4	7.1	0.5
American Indian or Alaska Native alone	0.7	0.2	0.3	0.3	0.7	0.2
Asian alone	3.7	0.3	5.3	1.5	3.6	0.3
Native Hawaiian or Pacific Islander alone	0.2	0.1	--	0.6	0.2	0.1
Some other race alone	2.7	0.4	2.2	1.1	2.7	0.4
Two or more races	2.2	0.2	3.1	1.2	2.2	0.2
Hispanic origin of householder	10.2	0.6	10.5	2.2	10.2	0.6
Percent of couples interracial	9.9	0.5	7.1	1.7	10.1	0.5
Average household income (in dollars)	114,476	1,452	83,505	7,616	116,500	2,127
Tenure						
Owns with a mortgage	50.6	0.8	39.4	3.1	51.3	0.8
Owns free and clear	26.2	0.7	40.3	2.8	25.3	0.7
Rents	23.2	0.7	20.3	2.6	23.4	0.7

-- Represents that the estimate or standard error is zero or rounds to zero

[1] Missing data could includes those with relationship and/or sex missing or had their data allocated for some reason, which includes "Don't Know" or "Refused."

Note: Unless otherwise specified, data in this table are percentages.

Source: U.S. Census Bureau, 2013 American Community Survey 1-year data file.

Table 3. Same-sex Married Couples by Whether Sex Assigned Based on the Names Index Matches Reported Sex by Response Mode

Response Mode	Total		Likely same-sex married		Not likely same-sex married		Cannot determine	
	Estimate	Std Error	Percent	Std Error	Percent	Std Error	Percent	Std Error
Total	251,695	4,169	67.7	0.8	15.8	0.6	16.5	0.7
Missing	15,440	998	55.8	3.1	8.9	1.5	35.3	2.9
No missing data	236,255	3,996	68.5	0.9	16.3	0.6	15.2	0.7
Mailout/mailback	87,867	2,282	54.3	1.2	26.1	0.9	19.6	1.0
CATI/CAPI	43,690	2,776	64.6	2.8	16.6	2.0	18.8	2.4
Internet	120,138	2,056	78.6	0.9	8.0	0.6	13.3	0.8

Source: U.S. Census Bureau, 2013 American Community Survey 1-year data file.

Table 4. Characteristics of Same-sex Married Couples by Whether Sex Assigned Based on the Names Index Matches Reported Sex

Characteristics	Likely same-sex married			Not likely same-sex married			Cannot determine		
	Total	Missing relationship and/or sex	No missing data	Total	Missing relationship and/or sex	No missing data	Total	Missing relationship and/or sex	No missing data
Total (number)	170,429	8,614	161,815	39,830	1,371	38,459	41,436	5,455	35,981
Average age of householder (years)	52.6	60.3	52.2	61.0	67.2	60.8	52.7	58.4	51.9
Average age of spouse (years)	51.4	60.0	50.9	59.6	65.7	59.3	51.6	56.7	50.8
Average household size (number)	2.7	2.6	2.7	2.9	2.1	3.0	2.9	2.7	2.9
Householder with at least a BA	51.8	25.2	53.2	23.3	--	24.1	39.2	24.7	41.4
Both spouses with at least a BA	35.4	9.8	36.7	12.1	--	12.6	26.1	14.0	28.0
Lives in the midwest	16.0	23.9	15.6	26.0	33.6	25.7	18.0	25.6	16.8
Race of householder									
White alone	85.9	76.6	86.4	83.4	87.5	83.3	70.7	70.1	70.8
Black alone	6.1	13.9	5.7	8.1	12.5	8.0	12.8	14.0	12.6
American Indian or Alaska Native alone	0.7	0.6	0.7	0.5	--	0.6	0.4	--	0.5
Asian alone	2.2	3.3	2.1	2.3	--	2.4	11.4	9.8	11.6
Native Hawaiian or Pacific Islander alone	--	--	--	0.4	--	0.4	0.6	--	0.7
Some other race alone	2.8	1.8	2.8	4.4	--	4.6	0.9	3.3	0.5
Two or more races	2.3	3.9	2.2	0.8	--	0.9	3.3	2.8	3.4
Hispanic origin of householder	10.1	10.4	10.1	14.0	3.4	14.3	7.0	12.4	6.2
Percent of couples interracial	10.6	8.5	10.7	6.0	1.9	6.2	10.9	6.2	11.6
Average household income (in dollars)	123,299	93,874	124,866	86,007	59,557	86,950	105,552	73,151	110,464
Tenure									
Owns with a mortgage	53.5	42.2	54.1	43.2	42.9	43.2	45.6	34.0	47.3
Owns free and clear	23.2	42.6	22.1	39.6	50.2	39.2	26.0	34.3	24.8
Rents	23.3	15.3	23.8	17.2	6.9	17.6	28.4	31.6	27.9

-- Represents that the estimate or standard error is zero or rounds to zero

[1] Missing data could includes those with relationship and/or sex missing or had their data allocated for some reason, which includes "Don't Know" or "Refused."

Note: Unless otherwise specified, data in this table are percentages.

Source: U.S. Census Bureau, 2013 American Community Survey 1-year data file.

Table 4a. Standard Errors for Characteristics of Same-sex Married Couples by Whether Sex Assigned Based on the Names Index Matches Reported Sex

Characteristics	Likely same-sex married			Not likely same-sex married			Cannot determine		
	Total	Missing relationship and/or sex	No missing data	Total	Missing relationship and/or sex	No missing data	Total	Missing relationship and/or sex	No missing data
Total (number)	3,402	743	3,286	1,533	245	1,498	1,945	554	1,887
Average age of householder (years)	0.3	1.4	0.3	0.6	2.0	0.6	0.7	2.0	0.7
Average age of spouse (years)	0.3	1.2	0.3	0.6	1.7	0.6	0.7	2.1	0.7
Average household size (number)	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1
Householder with at least a BA	1.2	4.1	1.2	1.7	6.3	1.7	2.2	4.8	2.2
Both spouses with at least a BA	1.1	2.5	1.1	1.5	6.3	1.5	2.0	3.7	2.1
Lives in the midwest	0.7	3.8	0.7	1.6	9.0	1.6	1.6	5.0	1.8
Race of householder									
White alone	0.8	3.9	0.9	1.3	6.2	1.4	1.9	4.7	2.2
Black alone	0.5	3.2	0.6	1.0	6.2	1.1	1.5	3.4	1.7
American Indian or Alaska Native alone	0.3	0.5	0.3	0.2	6.3	0.2	0.2	1.6	0.3
Asian alone	0.3	1.8	0.3	0.6	6.3	0.6	1.3	2.9	1.4
Native Hawaiian or Pacific Islander alone	--	1.1	--	0.3	6.3	0.3	0.4	1.6	0.5
Some other race alone	0.4	1.3	0.5	0.9	6.3	1.0	0.4	2.5	0.3
Two or more races	0.3	2.1	0.3	0.3	6.3	0.4	0.6	1.3	0.7
Hispanic origin of householder	0.7	2.8	0.7	1.3	3.4	1.4	1.1	4.0	1.1
Percent of couples interracial	0.6	2.5	0.6	0.9	1.9	0.9	1.3	2.5	1.4
Average household income (in dollars)	2,556	12,633	2,605	4,234	13,143	4,356	5,271	7,266	5,929
Tenure									
Owns with a mortgage	1.1	4.4	1.1	2.2	11.0	2.2	1.9	5.2	2.1
Owns free and clear	0.8	4.2	0.9	2.1	10.7	2.1	1.7	5.5	1.8
Rents	1.0	3.1	1.0	1.5	4.8	1.5	1.7	5.8	1.8

-- Represents that the estimate or standard error is zero or rounds to zero

[1] Missing data could includes those with relationship and/or sex missing or had their data allocated for some reason, which includes "Don't Know" or "Refused."

Note: Unless otherwise specified, data in this table are percentages.

Source: U.S. Census Bureau, 2013 American Community Survey 1-year data file.