

DRAFT

**Same Sex Households in the United States Census:
Measurement Issues and Substantive Results**

Howard Hogan, Martin O'Connell and Sarah Feliz

U.S. Bureau of the Census

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical and methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.

DRAFT

INTRODUCTION

The 2010 Census marks the first time that decennial census data will be shown for same-sex couple households by whether the couples reported themselves as living together as spouses or unmarried partners. Previously, data from Census 2000 combined both household types as there were no states in 2000 that performed same-sex marriage ceremonies. At the time of the 2010 Census, 5 states (Connecticut, Iowa, Massachusetts, New Hampshire, and Vermont) and the District of Columbia issued marriage certificates to same-sex couples. In addition, in May 2008, the California Supreme Court ruled that same-sex couples had a right to marry in California but that ruling was overturned by a ballot initiative in November 2008. There were also 3 states that did not perform same-sex marriages but recognized them from other states (Maryland, New York, and Rhode Island). By Census day in 2010, there were 37 states that had statutes defining marriage as only between one man and one woman while 10 states had laws prohibiting same-sex marriages.¹

Because of the relatively small number of same-sex marriages actually occurring in the United States (estimated at less than 50,000)² and the changes over the decade in state marriage laws, the data warrant careful analysis and evaluation. This technical paper will address the

¹ For a more complete overview of state laws, see the webpage of the National Conference of State Legislatures, <http://www.ncsl.org/default.aspx?tabid=16430>

² By 2010, there were estimated to be as many as 50,000 same-sex marriages performed in the United States and possibly up another 30,000 performed in other countries to U.S. residents. In addition, there were possibly up to another 85,000 same-sex couples in civil unions or domestic partnerships according to estimates prepared by researchers in the Williams Institute of the UCLA School of Law <http://www3.law.ucla.edu/williamsinstitute/pdf/Pressrelease2.24.pdf>

DRAFT

statistical aspects of the collection and presentation of data on same-sex couple households by the Census Bureau. We will first present overall trends from the 2000 and 2010 Censuses at the state and national level from tabulated data compiled from Summary File 1 representing the 100 percent counts of households. Next, comparisons between Census 2010 data and the 2010 American Community Survey (ACS) will be made to evaluate the consistency of the estimates between the two data sets.³ The third section will present techniques for evaluating the quality of the reports of same-sex couple data from the 2010 Census for both reports of same-sex unmarried partners and spouses. In light of this evaluation, the final section will offer alternative estimates of the same-sex couple household population from both the 2010 and 2000 Census.

STATE AND NATIONAL DATA FROM THE 2000 AND 2010 CENSUS

National Level Trends

According to Census 2010, the total number of same-sex couple households was 901,997, representing less than 1 percent (0.773 percent) of all households in the United States (Text Table A). Data from Census 2000 tabulated 594,391 households or 0.564 percent of all

³ Since 2005, statistics on characteristics of same-sex couple households—both for those identified as spouses and as unmarried partners—have been released from the ACS. The American Community Survey is nationally representative household survey of 3 million households that contains detailed social, economic, and housing items formerly contained on the long or sample form of the decennial Census. See the tables under the Same-Sex Couples heading at <http://www.census.gov/population/www/socdemo/hh-fam.html>

DRAFT

households. Overall, census data show an increase of 52 percent in the number (307,606) of same-sex households over the past ten years.⁴

**Text Table A. Same-sex Couple Households:
Census 2000 and 2010 Census**

| Household type | Number | | Percent change | Percent of all households | |
|---|-------------|-------------|----------------|---------------------------|-------|
| | 2010 | 2000 | 2000-2010 | 2010 | 2000 |
| Total households | 116,716,292 | 105,480,101 | 10.7 | 100.0 | 100.0 |
| Total same-sex couple households | 901,997 | 594,391 | 51.8 | 0.773 | 0.564 |
| Unmarried partners | 552,620 | 341,014 | 62.1 | 0.473 | 0.323 |
| Spouses | 349,377 | 253,377 | 37.9 | 0.299 | 0.240 |
| Percent of all same-sex couple households | 38.7 | 42.6 | (X) | (X) | (X) |

(X): Not applicable.

A greater percentage increase is observed among same-sex unmarried partner households (62 percent) than among same-sex spousal households (38 percent) since 2000. Unmarried partner households increased by 211,606 compared with an increase of 96,000 in spousal households. Overall, about 4 out of every 10 same-sex couple households both in 2000 and in 2010 were spousal households.

⁴ The data in this paper for Census 2000 represent the first time data for 2000 are shown for same-sex couple households separating the aggregate totals into unmarried partner households and spousal households for the Nation and individual states.

DRAFT

When examining these data, it should be noted that no state issued marriage licenses to same-sex couples in 2000. However, some households may have reported themselves as living together as spouses because they were in a civil union or domestic partnership, had made marriage-like commitments in a ceremony (although not sanctioned by state law), or had determined that this category best expressed their current household relationship. This count may have also included opposite-sex couples who were inadvertently included in the published Census 2000 tables as same-sex couples. This would occur if a partner in an opposite-couple household had their sex incorrectly marked or coded, making it appear that they were a same-sex couple. These data capture errors could include mismarks on the gender item on the Census form by either the respondent or the interviewer or errors in the processing phase when transferring data on the form to an electronic media.

By 2010, the population of same-sex spousal households also included those who were legally married either in the United States or foreign countries and were residents of the United States by the time of the April 2010 Census. It also may have included those couples in registered domestic partnerships or civil unions who found the relationship options limited on the form in describing their current living arrangements. A study of 602 same-sex couples in 2010 by the Williams Institute, UCLA School of Law, designed to understand how people completed their 2010 Census form and filled out the relationship item, found that 16 percent of couples who

DRAFT

were in a civil union or registered domestic partnership indicated that they selected the husband/wife relationship option.⁵

Geographic Patterns

Although same-sex couple households were less than one percent (0.773 percent) of households in the United States in 2010, distinct geographic patterns are evident. The pattern observed in Map 1a⁶ for all same-sex couple households shows states with above average percentages are found along the east and west coasts of the United States and in the southwestern part of the nation. When looking at state-level data for same-sex unmarried partner households, the geographic distribution for these households in Map 1b follows a similar pattern to the one found for all households (Map 1a). This is to be expected as 61 percent of same-sex couple households in 2010 were unmarried partner households.

A different pattern appears for same-sex spousal households. Map1c shows that a number of states reporting as same-sex spousal households with above average percentages are in the southern half of the nation ranging from Georgia to New Mexico. In addition, all of these states have explicit laws or state constitutions prohibiting same-sex marriages. None of these states have experienced legislative changes akin to those occurring in the east and west coasts of the

⁵ Gary Gates, "Same-sex couples in US Census Bureau Data: Who gets counted and why," The Williams Institute, UCLA School of Law, August 2010

http://services.law.ucla.edu/williamsinstitute/pdf/WhoGetsCounted_FORMATTED1.pdf

⁶ The ArcGIS maps in this report were created using 2010 Census source data. The data was classified using the natural breaks method, along with manual manipulation, where the national average was set to divide the upper and lower categories. The natural breaks method uses the Jenks Optimization mathematical algorithm, also known as the goodness of variance fit (GVF), when determining the value categories. More information can be found at <http://support.esri.com/en/knowledgebase/techarticles/detail/26442>.

DRAFT

United States in the past decade that would help explain the above average percentages of same-sex spousal households. This unexpected finding for these states will be addressed in a later section of this report.

The next series of Maps 2a-2c show the percentage increase in the number of same-sex households between 2000 and 2010 (52 percent, nationally). It is important to note that states with smaller numbers of initial same-sex household populations required less of a numerical increase to reflect an above average percentage growth than states, like California, with an already large number same-sex households. A pattern of high percentage increases for unmarried partner households in Map 2b is most notable for states ranging from the Carolina's through Oklahoma to the upper Midwest and the states in the Mountain region. Below average growth rates are found in California, Washington, and in the mid-Atlantic and New England areas. A somewhat different pattern is shown in Map 2c where the states with some of the highest growth rates in same-sex spousal households (40 percent or more) are Iowa, Vermont, Massachusetts, New Hampshire, Connecticut, California, and the District of Columbia (Appendix Table 2c). These relatively higher levels of increase are consistent with legislative changes that occurred in these areas in the past ten years.

Figure 1 illustrates the pattern of decadal increases in unmarried partner vs. spousal households among the states. The plot of these two indicators shows that states with low percentage increases in unmarried partner households had relatively higher percentage increase in spousal households, possibly the result of legislative changes regarding same-sex marriages. States that are outliers in the scatter plot (Massachusetts, Vermont, New

DRAFT

Hampshire, Connecticut, Iowa, California and District of Columbia) have all experienced various legislative changes in the past ten years concerning same-sex marriages. It is possible that the observed lower levels of increase in the percentage of unmarried partners in these outlier states is due to a shift to couples getting married over the decade instead of continuing to live as unmarried partners. This would simultaneously increase the spousal population while decreasing the unmarried partner population. States offering same-sex marriage opportunities may also prompt residents to form married households instead of unmarried partner households.

COMPARISONS WITH DATA FROM THE AMERICAN COMMUNITY SURVEY

The American Community Survey (ACS) is the Nation's primary annual source of intercensal population and housing data for sub-national geographical areas.⁷ The Census Bureau has published estimates of both same-sex couple unmarried partner and spousal households from this survey since it expanded its sample to 3 million households in 2005⁸. Figure 2 shows the overall national trend in same-sex couple households from the 2000 and 2010 decennial censuses and the intervening data from the ACS for 2005-2010.⁹

From 2005-2007, ACS estimates of same-sex couple households numbered between 750,000 and 800,000 households. A sharp drop was noted for the 2008-2010 rounds of the ACS to less than 600,000 households, with the decline in the post-2007 surveys principally occurring in the

⁷ For a more complete description of this survey, go to the Census Bureau's ACS webpage at <http://www.census.gov/acs>

⁸ See the "Same-sex Couples" section on the Families and Living Arrangements webpage at <http://www.census.gov/population/www/socdemo/hh-fam.html>

⁹ Same-sex couple data from the American Community survey for these years can be found in the Same-sex couples section in the Census Bureau's Families and Living Arrangements webpage <http://www.census.gov/population/www/socdemo/hh-fam.html>

DRAFT

same-sex spouse component from 341,000 in 2007 to about 150,000 households between 2008 and 2010. The 2010 Census, however, shows a sudden rebound in the total number of same sex-couple households to 902,000 with the majority of the change occurring in the same-sex spouse component to the previous levels (about 350,000) noted in the pre-2008 ACS surveys.

The 2007-2008 decline in the ACS has been attributed to both improvements in the editing and processing of the data and the layout of the mailout/mailback form, the latter change reducing data capture errors for the gender item which would produce “false positive” estimates of same-sex spouses from opposite-sex couples.¹⁰ The pre-2008 ACS used a “matrix” format which may be more prone to producing data capture errors in the gender item than in the current columnar arrangement mailout/mailback form (see Appendix Figures 1 and 2 for examples of sequential and matrix format designs for mailout/mailback forms). Because the errors would be generated from the population base of opposite-sex spouses (estimated to be about 55 million households after 2008), extremely small data capture errors of only 1 per 1,000 opposite-sex spousal households would generate 55,000 erroneously identified same-sex spouses in the tabulations. The number of opposite-sex unmarried partners, which were estimated at 5.6 million in 2008, would produce fewer same-sex unmarried partner households in error (5,600) at the same error level and would have a smaller impact on this estimate as there are many more same-sex unmarried partners than same-sex spouses.

¹⁰ See “New Estimates of Same-Sex Couple Households from the American Community Survey,” by Martin O’Connell, Daphne Lofquist, Tavia Simmons, and Terry Lugaila, Paper presented at the Annual Meetings of the Population Association of America, Dallas, Texas, April, 15-17, 2010.
http://www.census.gov/population/www/socdemo/hh-fam/SS_new-estimates.pdf

DRAFT

Figure 3 illustrates the discrepancies at the state level using data from the 2000 and 2010 censuses and the 2010 ACS. The state level numbers of same-sex couple households for the 2010 Census were greater for every state than in Census 2000. They were also greater than the 2010 ACS data for all states with the exception of Minnesota, North Dakota and Rhode Island (see Appendix Table 2a). What could account for this difference in same-sex couple households between these two data sources in 2010 by 52 percent? The likely answer for the difference between the 2010 ACS and the 2010 Census can be found in the prior explanation for the sudden drop in these numbers between the 2007 and 2008 ACS, namely, the forms being used for the follow-up phase for non-respondents to the mailout/mailback questionnaires.

Text Table B. Distribution of Households by Form Types: 2010 Census and 2010 ACS

| Form type | Census 2010 All households | Census 2010 same-sex households | Percent 2010 that are same-sex households | 2010 ACS All households* | 2010 ACS Same-sex households* | Percent 2010 ACS that are same-sex households |
|------------------------------|----------------------------|---------------------------------|---|--------------------------|-------------------------------|---|
| Total forms | 116,716,292 | 901,997 | 0.773 | 114,567,419 | 593,324 | 0.518 |
| Percent | 100.0 | 100.0 | | 100.0 | 100.0 | |
| Mail | 70.7 | 66.2 | 0.724 | 56.9 | 73.2 | 0.667 |
| Non response followup (NRFU) | 22.6 | 27.3 | 0.934 | NA | NA | NA |
| CATI/CAPI | NA | NA | NA | 43.1 | 26.8 | 0.322 |
| Other | 6.7 | 6.5 | 0.749 | NA | NA | NA |

* Weighted estimate.

DRAFT

For the both the 2010 Census and the rounds of the ACS beginning in 2008, the initial mailout/mailback phases of data collection used similar paper forms that collected individual household members' information in a sequential or vertically arranged format.¹¹ Among same-sex households, 66 percent of 2010 Census households and 73 percent of ACS households used the mail forms (Text Table B). For both of these data sources, about 0.7 percent of households using the mail forms indicated that they were same-sex households. However, the Census nonresponse follow-up (NRFU) forms returned a higher proportion of same-sex couple households (0.9 percent) than the ACS CATI/CAPI instruments (0.3 percent) which were responsible for the nonresponse follow-up phases of data collection.

The principal paper form used in the non-response follow-up phase (NRFU) phase of the 2010 Census was formatted similarly to the problematic ACS matrix form.¹² Comparisons between the Census and the ACS are further complicated because the ACS did not employ a paper form for non-response follow-up but collected information in person using a laptop computer (CAPI) or by telephone using a computerized instrument (CATI). These instruments had a key collection component absent from the 2010 Census NRFU paper form. When an ACS interviewer recorded that a household contained a spouse who was the same sex as the householder, a message in a "popup box" appeared on the instrument screen that prompted the interviewer to verify the responses to the relationship and the gender items. Previous

¹¹ See Appendix Figure 1 for the 2010 Census mailout/mailback form.

¹² Overall, there were 117 different form types used in the 2010 Census. Many of these forms were variants of other forms developed to handle different types of housing (for example dormitories, military bases, hospitals) and for specific statistical experiments. For purposes of this report and to attempt to provide forms that are similar in format and content to the ACS, 2010 Census forms designated as "mail" were form types 111-153 and "NRFU" were form types 401-423. See Appendix Figure 2 for the enumerator administered NRFU form. See Appendix Figure 2 for the matrix format design of the questionnaire.

DRAFT

Census Bureau research has demonstrated the sensitivity of same-sex couple estimates to these different data collection formats.¹³

Text Table B indicates that 27 percent of all same-sex couple households in the Census 2010 had non-response follow-up interviews conducted by Census Bureau interviewers using the matrix formatted paper form. Similarly, non-response follow-ups made up 27 percent of all same-sex households in the ACS but were conducted using either CATI or CAPI interviews. Map 3a shows a pronounced regional variation in the use of NRFU forms in the 2010 Census which shadows the 2010 Census mail participation rates as nonparticipation resulted in non-response follow-up.¹⁴ States in the southern section of the United States and in New England had above average portions of households using NRFU forms while the Midwestern states were generally below average. Looking only at same-sex couple households, Map 3b shows that NRFU forms were used in above average proportions for states running from the southern Gulf coast to the upper northwest.

Figure 4 graphs the differences in the percentage of households reporting as same-sex couple households by form types for the 2010 Census and the 2010 American Community Survey. Overall, about 0.8 percent of households in the 2010 Census were identified as same-sex couple households compared with 0.5 percent in the 2010 ACS. The percentages for those using the similarly formatted mailout/mailback forms averaged 0.7 percent each, although statistically higher in the 2010 Census than in the 2010 ACS (0.724 percent and 0.667 percent,

¹³ O'Connell et al., op. cit., 2010.

¹⁴ See the Census webpage on participation rates for the 2010 Census <http://2010.census.gov/2010census/take10map/>

DRAFT

respectively)¹⁵. However, among the non-mail returns¹⁶, the 2010 Census reported same-sex couple households at a level three times that reported by the 2010 ACS (0.9 percent and 0.3 percent, respectively).

Higher percentages of same-sex couple households are noted for every state in the 2010 Census compared with the 2010 ACS (Figure 5a) with the exception of 5 states (Appendix Table 3). However, differences are evident between the two data sets when analyzed by the mode of data capture. Only 14 states had higher percentages in the 2010 Census than in the 2010 ACS among those using the mail form (Figure 5b) but among those using the non-mail forms (Figure 5c), 48 states had higher levels in the 2010 Census (see Appendix Table 3 for the percentages and standard errors of the ACS estimates). Summarizing the differences in the mode of data capture for the 2010 Census, Figure 6 clearly demonstrates that the NRFU forms recorded higher percentages of same-sex couple households of all households for every state with the exception of California, Washington, and the District of Columbia.

This pattern of higher rates of same-sex households in the 2010 Census than in the 2010 ACS and by form type within the 2010 Census demands a further investigation into the quality of the 2010 Census data. The following sections will provide an indirect way of evaluating the data captured and to offer a way to estimate the number of same-sex households from the 2010 Census that accounts for potential data capture errors.¹⁷

¹⁵ All levels of differences in this report are stated at the 90 percent level of confidence.

¹⁶ In this instance and for the remainder of the report, the phrase “non-mail” forms indicates the previously listed NRFU forms for the 2010 Census and the CATI/CAPI instruments for the ACS follow-up interviews.

¹⁷ Data capture errors may include transcription mis-markings by respondents and enumerators and also errors in data processing by optical equipment used in scanning the forms.

DRAFT

DATA QUALITY ISSUES OF SAME-SEX COUPLE HOUSEHOLDS IN THE 2010 CENSUS

Measurement technique

The classification of households relies both on the accuracy of the responses to the household relationship item—either as a spouse or unmarried partner—and to the gender item. Previous analysis of responses to the gender item in Census 2000 shows that this item had both the lowest allocation rate (0.9 percent) and index of inconsistency (1.7 percent) of all items on both the short and long forms.¹⁸ However, the first names of respondents may occasionally be at odds with their reports on gender. Because the number of same-sex couple households is relatively small, minor errors in the gender item made by opposite-sex couple households could have a substantial impact on the estimates of same-sex couple households.¹⁹

Without having information from a detailed content re-interview concerning the accuracy of the joint reporting of the gender and relationship items for same-sex couple households from the 2010 Census, an indirect method of evaluation is offered using the consistency of the first names reported on the forms and the data recorded for the gender item. This provides us with a way to measure inconsistencies in the reporting of the gender item that affects the numbers

¹⁸ The index of inconsistency is a measure of response variance in questions. Part of the Census 2000 program was to conduct a content re-interview Survey to measure the consistency of responses between questions on Census 2000 and a subsequently administered survey. For a description of this survey and the ensuing analysis, see Paula J. Schneider, *Content and Data Quality in Census 2000*, Census 2000 Testing, Experimentation, and Evaluation Program Topic Report No. 12, TR-12 (US Census Bureau: Washington DC, 2004), Table 1.

¹⁹ The likelihood of data capture errors for the relationship item inflating the number of same-sex households will be discussed in Appendix A.

DRAFT

of same-sex couple households, and also a way to produce estimates of same-sex couple households that address the likelihood of an error being committed.

An earlier paper using the name-sex inconsistency technique was developed from test census data collected in New York in 2004.²⁰ The Census Bureau has developed statistical “name directories” which are files of first names that are associated with a probability index that identifies the “maleness” of the name. These name directories for the 2010 Census were developed for each state from internal micro-data from individual-level data records. The probability index (from 0 to 1000) for each name in the directory was constructed by taking the ratio of the number of times this name was recorded with a male response to the gender item to the total number of times this name was recorded either as male or female.

For example, an index of 950 indicates that when this name appeared in the census for a given state, 950 times out of 1000, that person was a male. An index of 20 would indicate that only 20 times out of 1000 that name was reported with a male response or conversely, 980 times out of 1000 that name was identified as being reported female. A decision, then, could be made as to whether to accept the person’s report of their sex on the basis of consistent reports with this index or to reject their response on the gender item. Clearly, age, cultural and geographical differences may affect this probability, as similarly spelled names may be male or female in different cultural environments. Census directories prepared at the state level partly address these issues.

²⁰ See Martin O’Connell and Gretchen Gooding, “Editing Unmarried Couples in Census Bureau Data,” July 2007. <http://www.census.gov/population/www/documentation/twps07/twps07.html>

DRAFT

By setting different “acceptance levels” for this index, one can see the effect of using an alternative piece of information—a person’s first name—in the review or editing of data files.²¹ For example, suppose one was confident that an error was made in marking the gender item as “male” if a person’s name only 5 percent of the time was recorded as “male” in the names directory (meaning that 95 percent of the time that name was associated with reports by female respondents). A name more likely to have both male and female responses (for example, Leslie compared with John) would have a lower index level. A decision to reject the sex response for names with lower index values would have a greater potential for making an error when the response to the gender item was actually correct.

Text Table C illustrates how the first name index can be used to identify data capture errors in the gender item using data from the 2010 Census, using, as an example, male-male same-sex couple households in Texas.²² Ideally, one would expect a large proportion of all male-male households with both the householder and the partner having names with high level indices, indicating that their name is likely to be associated with being male. Respondents with values of 0-50 indicate that their name is associated with being a female for 95 percent or more of the cases where name and sex are recorded on the census forms.

²¹ Some 2010 Census editing routines did use a person’s name to assign a male/female value for the gender item when that question was left blank on the form and no other useful information was available for editing procedures.

²² See Appendix Table B for the corresponding female-female first name table for Texas.

DRAFT

Text Table C. Illustrative Example: First Name Indices for Male-Male Households in Texas, 2010 Census
(Numbers of reported partners)

Partner Name Index

| Householder Name Index | 0-50 | 51-100 | 101-499 | 500-899 | 900-949 | 950-1000 | NA | Total |
|------------------------|-------|--------|---------|---------|---------|----------|-------|--------|
| 0-50 | 35 | 4 | 36 | 90 | 54 | 1869 | 105 | 2,193 |
| 51-100 | 13 | 0 | 2 | 7 | 1 | 77 | 6 | 106 |
| 101-499 | 125 | 8 | 51 | 34 | 11 | 361 | 32 | 622 |
| 500-899 | 302 | 22 | 119 | 132 | 15 | 621 | 88 | 1,299 |
| 900-949 | 178 | 12 | 24 | 27 | 10 | 339 | 28 | 618 |
| 950-1000 | 6,265 | 261 | 699 | 786 | 411 | 14,439 | 1,184 | 24,045 |
| NA | 275 | 24 | 58 | 53 | 32 | 498 | 1,942 | 2,880 |
| Total | 7,193 | 331 | 987 | 1129 | 534 | 18,204 | 3,385 | 31,763 |

Of the 31,763 male-male households in Texas, less than one half (14,439 in green) had both partners' names having a name index of between 950-1000, indicating that 95 percent or more of the time their first names indicated a male response to the gender item. For 2,193 households, the "male" householder had a name with an index value of 50 or less and for another 7,193 households, the partner had a name with an index value of 50 or less, strongly suggesting that these respondents were actually female (numbers in red). Including the 35 households where both partners indicated a gender inconsistent with their first name²³ there were 9,351 male-male households in Texas (29 percent of all male- same-sex households)

²³ These cases were apparently so inconsistent and suspect that they were treated as not having a reasonable set of responses that could be considered as correctly identifying a same-sex household. These double inconsistencies generally represented about 1 in 1,000 of all same-sex couple households for all states.

DRAFT

which had at least one partner with a female name, therefore, erroneously inflating the number of male same sex-households.

For this exercise, a level of 0-50 was selected to indicate that a data capture error had occurred; similarly, a level of 950-1000 was selected for female-female households indicating that a female householder or partner was likely to be male but that a data capture error was made for the gender item. This conservative approach was selected to minimize the problem of having a name with an ambiguous gender (for example, Morgan, Kelly, Jean) that would be considered to be in error with the gender response while in fact the gender response was correct.²⁴

Increasing the index level to 100 would only add another 437 cases out of 31,763 households in the present example.

To estimate the level of potential data capture errors among same-sex unmarried partners, an inconsistency ratio was constructed showing the number of same-sex unmarried partners with inconsistent name and sex reports per 100 opposite-sex unmarried partner households. For same-sex spouses, the ratio was the number of same-sex spouses with inconsistent name and sex reports per 100 opposite-sex spouses. The denominators of both ratios (opposite-sex partners and opposite-sex spouses) are the sources of these data capture errors. In no case was any same-sex spouse household classified as such if either the gender or the relationship item for either the householder or spouse had been changed or imputed in any way in the editing process.

²⁴ Names with no index value are those with less than 10 occurrences in a state, where the first name is left blank or has only character, or where no one in that state with that name reported their gender.

DRAFT

Text Table D shows the level of inconsistency between name-sex reporting for the two types of living arrangements (partners and spouses) for the mail and NRFU forms from the 2010 Census. The states that were selected to illustrate this table had varying degrees of recognition of legal statuses of same-sex couples at the time of the 2010 Census, from allowing same-sex marriages (Connecticut, Massachusetts, California, and the District of Columbia) to allowing same-sex civil unions (New Jersey) to neither allowing nor recognizing same-sex marriages (Oklahoma and Texas).²⁵

Text Table D. Inconsistent Name-Sex Reporting Per 100 Opposite-sex Unmarried Partners and of Opposite-sex Spouses: Selected States, 2010 Census

| Inconsistent reporting of names and sex resulting in same-sex couples | | | | |
|---|--|------|--|------|
| State | Inconsistent SS unmarried partners per 100 OS unmarried partners | | Inconsistent SS spouses Per 100 OS spouses | |
| | Mail | NRFU | Mail | NRFU |
| Total, United States | 0.34 | 1.10 | 0.24 | 1.05 |
| Connecticut | 0.33 | 0.90 | 0.26 | 0.95 |
| Dist of Columbia | 0.44 | 0.96 | 0.31 | 1.10 |
| Massachusetts | 0.36 | 1.17 | 0.27 | 1.05 |
| California | 0.37 | 0.99 | 0.24 | 0.89 |
| New Jersey | 0.37 | 1.23 | 0.25 | 0.96 |
| Oklahoma | 0.36 | 1.16 | 0.24 | 1.18 |
| Texas | 0.40 | 1.14 | 0.24 | 1.04 |

The first pattern noticed is that the name-sex inconsistency ratio at the national level and for the states shown was 3-4 times greater for the NRFU forms than for the mail forms for both unmarried partners and spouses: about 1.1 per 100 for the NRFU forms and 0.2-0.3 per 100 for the mail forms. The second pattern is that within each form type, the inconsistency ratios

²⁵ Appendix Table 4 contains these name-sex inconsistency ratios for all states.

DRAFT

among the states are within a narrow range regardless of whether the data capture error is being measured for partners or spouses.

This is shown in the box plot (Figure 7) which illustrates both the level of the inconsistency ratios and how they are clustered by household type (unmarried partners v. spouses) and form type (mail v. NRFU) for the 50 states and the District of Columbia. The horizontal caps at the end of each vertical line indicates the highest and lowest ratio recorded excluding those areas that are more than 3 standard deviations from the median—if there are any such states, they are labeled and positioned accordingly on the chart. The second (dark blue) and third quartile (light blue) show the distribution around the median.

The level of the inconsistency ratios strongly suggests that the data capture error is primarily form based and not based in the type of household being interviewed, although opposite-sex partners seem to generate slightly higher levels of name-sex inconsistency ratios (0.34 per 100) than opposite-sex spouses (0.24 per 100). The clustering of the level of these errors among the states around the median further suggest that there is something inherent in the construction of the NRFU questionnaire format (the matrix approach) that is generating a higher degree of inconsistent reports between the name and sex of the respondent than in the mail forms.²⁶

How Do Data Capture Errors Affect the Estimates of Same-sex Couple Households?

²⁶ For a detailed analysis of the potential problems associated with collecting these data in matrix format compared with the vertically formatted mail form, see Kathleen T. Ashenfelter, “Eye-tracking Study Report: Examining User Patterns for Demographic Items on the 2007 and 2008 Mail Forms,” Survey Methodology Study Series (SSM 2010-01), Statistical Research Division, U.S. Census Bureau, January, 2010 <http://www.census.gov/srd/papers/pdf/ssm2010-01.pdf>

DRAFT

While the magnitude of these inconsistency ratios may seem trivial (less than 1 percent), it is important to understand that their effect is greatly magnified when applied to extremely large population bases—56.5 million opposite-sex spouses and 6.8 million opposite-sex unmarried partners in Census 2010. Two series of numbers for same-sex couple households from the 2010 Census (and later from Census 2000) will be discussed in the following sections of this report. The first set of numbers will be those without any name-sex inconsistency corrections and are those currently shown in the American FactFinder when detailed tables from Summary File 1 are displayed. These counts of same-sex couple households from the 2010 Census will be noted as “Summary file counts.” The second set of estimates of same-sex couples contain the name-sex inconsistency corrections and will be noted as the “Preferred estimates” as they remove the likely numbers of couples that are incorrectly included in the same-sex couple households and are erroneously inflating the estimated population. These estimates are the recommended set of numbers to be used when examining both levels and relative differences in the number of same-sex couple households.

Figure 8 shows the impact of these data capture errors on the number of same-sex households from the 2010 Summary file counts. Overall, 28.3 percent of same-sex couple households in 2010 are likely to be opposite-sex households (individual state data are shown in Appendix Tables 5a and 6b). This would drop the number of same-sex couple households in the 2010 Census presented from Summary File 1 on the American FactFinder from 901,997 households to an estimate of 646,464 households. About one-half of the same-sex couple households from the NRFU forms were likely to have data capture errors compared with 20 percent of those same-sex households from the mail forms.

DRAFT

As previously noted (Text Table D), the form based inconsistency ratios are of similar magnitude for both unmarried partners and spouses. However, because there are so many more opposite-sex spouses than partners to potentially generate data capture errors, the effect of these data capture errors for the spousal component is exaggerated when computing these overall proportions. Figure 8 shows that 7 percent of unmarried partners are estimated to have data capture errors compared with 62 percent of same-sex spouses.

For unmarried partners, the degree of inconsistent names and sex was almost 6 times as great for household data obtained from NRFU forms as from mail forms (22 percent and 4 percent, respectively). The impact of the data capture errors on names-sex inconsistency was greater for spouses (62 percent); 69 percent for non-mail NRFU forms and 57 percent for mail forms. As there were relatively few married same-sex couples in the United States in 2010, same-sex spouses derived from data capture errors from either the NRFU or mail forms greatly outnumbered the spousal responses not recorded in error and resulted in large proportions of falsely classified same-sex spousal households for both forms.

By using a box plot graphic, Figure 9 shows the variation in the percent of same-sex couples with inconsistent reporting of names and sex by type of living arrangement and the form used to obtain the information. For unmarried partners, the overall impact of data capture errors was greater for the estimates from NRFU forms than from mail forms. Relatively high percentages of inconsistent reporting for unmarried partners among mail returns were found for South Dakota, North Dakota, Wyoming and Mississippi.

DRAFT

The proportion of spousal households with inconsistent name-sex reporting was higher than that of unmarried partners for both mail and NRFU forms. However, relatively low proportions of inconsistency in mail forms were found for the spousal component for states in 2010 that had passed same-sex marriage laws by the time of the 2010 Census--Vermont, California, Massachusetts, and the District of Columbia. The latter two areas also had relatively low percentages for the NRFU returns. Relatively low percentages for these states can be attributed to relatively high numbers of correctly reported same-sex marriages in these areas (as they allowed marriages) which would minimize the overall impact of data capture errors on the total estimate.

State differences in the percent of same-sex couple households with inconsistent reporting of names and sex are shown in Map 4. Above average percentages in the upper part of the distribution are clustered in the south (Arkansas, Mississippi and Alabama) and the northern tier of states including North Dakota and South Dakota, Montana and Wyoming. Figure 10 plots the percent of same-sex couples using the NRFU forms and the percent of same-sex couples with inconsistent first names and reported sex. The graph clearly shows the relationship between the consistency of the relationship and gender responses obtained for same-sex couples and the usage of the NRFU form. With a strong positive correlation between these two indicators of 0.837, every 1.0 percentage point increase in the use of NRFU forms corresponded to a 1.1 percentage point increase in percentage of these couples having inconsistent reports between one of the partners' names and their sex.²⁷

²⁷ This percentage point association is determined by a regression fit of the NRFU percentages against the percentage of same-sex couples with inconsistent names and sex responses.

DRAFT

These data strongly indicate that data capture errors by opposite-sex couple households answering the NRFU forms created inflated estimates of same-sex couples. These erroneous counts are most notable in the states extending from the East South Central part of the United States through Dakotas and the Mountain states, all states which have either legislative or state constitutional language defining marriage as between only one man and one woman. The following section will re-examine state-level estimates of same-sex couples after removing from the 2010 Census summary file counts those households which have inconsistent name and sex responses.

ESTIMATES OF SAME-SEX COUPLE HOUSEHOLDS

Preferred Estimates for the 2010 Census

As shown in the previous section, potential misreporting of gender varied by geographic area, by data capture mode, and by whether the same-sex couple household was either an unmarried partner or a spousal household. What is the effect of all these factors on the 2010 Census numbers of same-sex couple households?

DRAFT

Text Table E. Same-sex Couple Households, Summary File Counts and Preferred Estimates: 2010 Census and 2010 ACS

| Household type | 2010 Census summary file counts | 2010 Census preferred estimates | 2010 ACS estimates* |
|--|---------------------------------|---------------------------------|---------------------|
| Total same-sex couple households | 901,997 | 646,464 | 593,324 |
| Unmarried partners | 552,620 | 514,735 | 440,989 |
| Spouses | 349,377 | 131,729 | 152,335 |
| Percent spousal households of all same-sex couple households | 38.7 | 20.4 | 25.7 |

* Weighted estimate.

Text Table E shows the different counts and estimates of same-sex couple households for the 2010 Census and the 2010 ACS. The 2010 preferred estimates remove from the summary file counts those couples where the names of the respondents are inconsistent with their reported sex at an index level of 95 percent or more, strongly suggesting that they are opposite-sex couples. Overall, the total number of same-sex couples declines from 901,997 to 646,464 or 28 percent. The unmarried partner component declines 7 percent while the spousal component declines 62 percent to a level (131,729 households) actually less than the 2010 ACS estimate (152,335 households with a standard error of 3,225).²⁸ The relative percentage of spousal households of all same-sex couple households declines from 39 percent to 20 percent, a proportion now less than the 2010 ACS estimate of 26 percent.

²⁸ It should be noted that the 2010 ACS estimates may also contain data capture errors generated by the mail forms that have not been removed from the data as variables needed to do so are not available from the data files.

DRAFT

Examining the percentage of all households that are same-sex households, Figure 11a illustrates that the 2010 Census preferred estimates are now closer but still higher than the ACS estimates, declining from 0.773 percent of all households to 0.554 percent compared with the ACS estimate of 0.518 percent (standard error of 0.006 percent). These declines are especially notable for spousal households from 0.299 percent to 0.113 percent. Figure 11b also notes that the declines are most notable for the NRFU forms, falling in half from 0.934 percent to 0.465 percent, compared with the ACS estimate of 0.322 percent (standard error of 0.009 percent). The decline in the mail form using the preferred estimate (0.580) is such that it is now less than the ACS estimate (0.667, standard error of 0.013). It should be noted that the ACS estimate for the mail form may also include false positives as was found in the Census mail form so in effect the ACS estimate may also be higher than actually shown and may in fact be less than the 0.667 reported in the survey.

Figures 12a-12c illustrate the changes in state-level 2010 Census and 2010 ACS comparisons using the 2010 Census preferred estimates--all three charts show that the preferred 2010 estimates are below the 2010 Summary file counts. Figure 12a shows that the percentages derived from the 2010 summary file counts were higher than the ACS estimates in 46 states--the preferred estimates now bring that number down to 13 (Appendix Table 6a). The improvements in Figure 12c for same-sex spouses are very noticeable as one would surmise from the aggregate percentages shown in Figure 11a using the preferred estimates.

DRAFT

The next series of Maps 5a-5c show the percent of same-sex couple households (by type) of all households using the preferred estimates. They can be compared with Maps 1a-1c to examine any differences from the statistics tabulated from the 2010 Census summary file counts.

The overall pattern of percentages shown in Map 5a, although at a lower level, is consistent with the geographic patterns shown Map 1a. Every state shown in Map 1a that had above average percentages of same-sex couple households also had above average percentages using the preferred estimates in Map 5a. Same-sex couple households are still more frequently found on the East and West coasts than in the central part of the United States. Map 5b also shows that the unmarried partner pattern is also similar in pattern to Map 1b.²⁹ This is an important finding as it supports the uses of the detailed tabulated 2010 Census data when making relative state-level comparisons for the total and unmarried partner tabulations.

However, Map 5c serves as a warning to data users when examining data on same-sex spouses. As previously noted, almost two-thirds of all same-sex spouses are potentially opposite-sex spouses that have been created by data capture errors. Map 1c from 2010 Census summary file counts shows a band of states with above average percentages of same-sex couple households extending from Georgia on the east coast through New Mexico. Map 5c shows a different pattern of same-sex spousal households using the preferred estimates.

As shown in previous sections, higher utilization of NRFU forms were found among same-sex couples in the South (Map 3b) and these forms were associated with a higher incidence of data-

²⁹ Appendix Tables 6a and 6b contain the estimates for the series of 5a-5c maps showing statistics from 2010 Census summary file counts, 2010 Census preferred estimates and 2010 ACS estimates for individual states and by type of same-sex couple household.

DRAFT

capture errors. While above average percents of same-sex spousal households are still found on the west coast and in the Mid-Atlantic and the New England areas, the above average percents formerly found in the South in Map 1c all but disappear in Map 5c when using the preferred estimates. The states from Georgia to New Mexico now all have percentages below the national average of 0.113 percent of all households.

With the issuance of 2010 Census data that separates same-sex couple households into unmarried partners and spouses, the data user will be able to obtain a better picture of geographic variations in same-sex couples by separating the more error prone data for same-sex spouses from the aggregate totals.

Preferred Estimates for the 2000 Census

The remaining issue discussed in this paper examines changes between 2000 and 2010 in same-sex couple households using an indirect method of estimating same-sex couples in Census 2000 generated by errors from opposite-sex couple households. Unlike the 2010 estimates, the preferred estimates for 2000 were not generated from the Census 2000 micro-data level files using the respondent's specific name and sex responses. The 2000 estimates were indirectly produced utilizing state-specific name-sex inconsistency ratios derived from the 2010 Census. Inconsistency ratios for each state for 2010 were developed for unmarried partners and spouses (Appendix Table 7). These 2010 inconsistency ratios were applied to the Census 2000 numbers of opposite-sex spouses and opposite-sex unmarried partners to produce the likely number of same-sex spouses and unmarried partners in 2000 who would be incorrectly

DRAFT

included in the same-sex population. These estimates of households were subtracted from the overall Census 2000 counts to produce similar preferred estimates for Census 2000.

Because of the narrow range of inconsistency ratios exhibited across states for similar forms (Figure 7), a critical assumption in this indirect method is that the data capture problems are form specific and not would vary between the 2000 and the 2010 Census. Census 2000 used both mail and NRFU paper questionnaires that were in the same format design as in the 2010 questionnaires and the mail participation rates for 2000 were of the approximate magnitude as in 2010 (74 percent) and hence would likely have the approximate same mix of mail and NRFU forms.³⁰ The national results are shown below in Text Table F while detailed state preferred estimates are presented in Appendix Table 7 for same-sex unmarried partners and spouses as well as for all same-sex couple households.

³⁰ The 2000 Census mail forms consisted of both short forms—like the 2010 Census—and long forms containing many more social, economic and housing items. However, both forms used the vertical/sequential alignment of questions instead of the matrix format.

<http://www.census.gov/dmd/www/2000quest.html>

DRAFT

Text Table F. Same-sex Couple Households, Summary File Counts and Preferred Estimates: 2000 and 2010 Census

| Household type | Summary file counts ¹ | | Summary file percent change | Preferred estimates ² | | Preferred estimate percent change |
|--------------------|----------------------------------|-------------|-----------------------------|----------------------------------|-------------|-----------------------------------|
| | Census 2010 | Census 2000 | 2000-2010 | Census 2010 | Census 2000 | 2000-2010 |
| Total | 901,997 | 594,391 | 51.8 | 646,464 | 358,390 | 80.4 |
| Unmarried partners | 552,620 | 341,014 | 62.1 | 514,735 | 314,052 | 63.9 |
| Spouses | 349,377 | 253,377 | 37.9 | 131,729 | 44,338 | 197.1 |

¹ Tabulated from internal 2000 and 2010 Census Summary Files.
² Preferred estimates for 2000 indirectly derived from 2010 Census inconsistency ratios and numbers from Census 2000 and 2010 Census, Appendix Tables 6b and 7.

From Text Table F, we see that in both 2000 and 2010, the preferred estimates for the unmarried partner population are much closer to the summary file counts than are the spousal numbers. In fact, the growth in the same-sex unmarried population between 2000 and 2010 is 62 percent using the summary file counts and 64 percent using the preferred estimates. This gives the data user confidence in the trends observed for same-sex unmarried partners over the decade using the summary file statistics. However, the spousal component presents some problems.

In 2000, no same-sex marriages were issued in the United States. Although no legal state-sanctioned marriages were performed in 2000, some couples may have had religious or church ceremonies that served as affirmation of a marriage but were not recognized by the laws of their state of residence. This leaves any counts of same-sex spouses in 2000 due either to data capture errors or to accurate reports by same-sex couples who felt that their relationship is

DRAFT

best described as living together as spouses. Because of this, elimination of the data capture error leaves a relatively small number of couples providing same-sex spousal reports in 2000.

As one can see, the data capture error component drops the count of same-sex spouses from 253,377 to 44,338, indicating that about 83 percent of all same-sex spouses in the 2000 Census were potentially opposite-sex married couples with a data capture error in the gender item.

Figure 13 illustrates state profiles of the relative size of the data capture error component (red line) compared with Census 2000 summary file counts (Appendix Table 7).

However, in 2010, there were areas that had legalized same-sex marriage or recognized marriages performed in other countries. The impact of data capture errors in 2010 had a significant but smaller effect as the population of same-sex spouses in 2010 included both those with data capture errors and socially defined (but not state sanctioned) marriages as in 2000, but additional numbers of state-sanctioned marriages. This third component proportionately reduces the overall impact of data capture errors. The greater data capture error component in 2000 (83 percent) than in 2010 (62 percent) actually makes the increase in the same-sex spouse component larger using the preferred estimates (197 percent) than the summary file counts (38 percent) which primarily consist of name-sex inconsistent households.

Relative decadal increases in the total number of same-sex couple households by state computed using the preferred estimates closely track those produced from the summary file counts. The correlation of these decadal increases for the two data series for the 50 states and the District of Columbia is 0.704 for all same-sex households and an extremely high 0.953 for

DRAFT

unmarried partner households.³¹ However, the correlation of the decadal increase among the states for the spousal households is very low (0.085) due to the high level of spousal households in 2000 that were probably represented by data capture errors.³²

From an analytical point of view, the *tabulated* data shown in either the detailed tables from Summary File 1 on the American Factfinder or in the current report that best represents the relative increases in same-sex households between the two censuses are the data for the unmarried partner component. It closely tracks the changes and patterns observed when using the preferred estimates prepared for this report. Even using the preferred series, the data presented in 2010 Census tables for same-sex couples by presence of children in the households should still be treated with caution. As opposite-sex couples are more likely to have children of their own than same-sex couples, a high proportion of incorrectly assigned same-sex couples may affect the accuracy of the data, especially among households that contain spouses.³³

SUMMARY

This report examines the problems when attempting to estimate relatively small populations from either Census or survey data that are derived from responses to multiple questions. In

³¹ See Appendix Tables 6b and 7 for the basic data to derive these correlation coefficients.

³² There were 6 states in 2000 that had the predicted errors exceeding the tabulated numbers: Iowa, Montana, New Hampshire, South Dakota, West Virginia and Wisconsin. These states were deleted from the calculation of the correlations for the spousal households.

³³ See data from the ACS on characteristics of opposite-sex and same-sex households on the Census Bureau's Families and Living Arrangements webpage regarding the presence of children in households.

<http://www.census.gov/population/www/socdemo/hh-fam.html>

DRAFT

this case, the estimation of the population of same-sex couple households is not derived from a single question about household living arrangements but is dependent on the accurate reports of *both* same-sex and opposite-sex couples to the questionnaire items on relationship to householder and the sex of the respondent. Since the number of same-sex couple households is under 1,000,000—and the number of same-sex spouses is under 50,000 in terms of legal marriages—the numbers reflected in either censuses or surveys may be highly influenced by the degree of data capture errors.

A comparison of 2010 Census data with the 2010 American Community Survey suggested that the design of the non-response follow-up forms in the 2010 Census could be the source of the inconsistency between the two data sets. This format was also used in the pre-2008 ACS and has been shown to potentially cause data capture errors which could inflate the numbers of same-sex couple households due to marking or transcription errors in the gender item.

Using comparisons of first names with sex responses on the Census forms, a method was developed to provide a preferred set of estimates of the number of same-sex couple households correcting for the inadvertent inclusion of opposite-sex. This paper has indicated that 28 percent of all same-sex couple households in the 2010 Census are likely to be opposite-sex couple households, thus inflating the numbers shown in Census Bureau products previously based on Summary File 1 counts and released between June-August 2011. This problem is more severe for those couples who reported being spouses (62 percent) than unmarried partners (7 percent). This problem can be traced to the format of the paper questionnaire

DRAFT

being used in the non-response follow-up (NRFU) phase of data collection. Approximately 10 same-sex couple households were incorrectly generated by data capture errors for every 1,000 opposite-sex households interviewed with NRFU forms compared with approximately 3 for every 1,000 opposite-sex households from mail forms.

As the ratio of opposite-sex to same-sex spousal households is much greater than the ratio of opposite-sex to same-sex unmarried partner households, the effects of these errors are disproportionately greater on the numbers of same-sex spouses than unmarried partners.

Since the proportion of same-sex households from NRFU forms is higher in the South and the upper Midwest, data from these areas are more affected with inflated counts than areas on the West coast and in Northeast.

Similar preferred estimates for same-sex couple households were indirectly constructed for Census 2000. Based on the assumptions that data capture error rates for 2000 were the same as in 2010, there was a strong correspondence in the level of the 10-year change between the original summary file counts and the preferred estimates for same-sex unmarried partners.

However, since a greater component of the 2000 data for same-sex spouses was attributable to data capture errors than for the 2010 data, the data suggest an even greater decadal increase in the spousal component when the preferred estimates are used. This stands to reason as in 2000 there were no same-sex marriages that were performed in the United States but by 2010 there may have been as many as 50,000 marriages performed in the Nation and up to 30,000 more to U.S. residents that were performed in other countries.

DRAFT

For analysis at national and state levels for 2010, researchers should use the preferred estimates presented in this report and in the additional table accompanying the release of this report showing the type of same-sex household by the sex of the partners and the presence of children in the household. For changes between Census 2000 and 2010, use Appendix Tables 6b and 7 in this report for obtaining preferred estimates that can be used to examine ten-year changes for the nation and for individual states.

There is no plan to retabulate a set of preferred estimates for the 2010 Census tables currently on the American Factfinder (PCT15 and PCT17) for sub-state geographical areas using the method outlined in this paper. However, the 2010 Census Public Use Microdata Sample (PUMS) will be a 10 percent sample of households and contain a variable on the file indicating if a same-sex partner household originally reported themselves as spouses and subsequently was modified to a same-sex unmarried partner household on the edited data file. This will enable data users to selectively examine the two different types of same-sex households at lower levels of geography than presented in this paper although the names index used in this report will not be available on the PUMS file. This file will provide geographical identifiers for households in Public Use Microdata Areas (PUMS) of 100,000 or more population.³⁴ As suggested in this paper, the estimates for the unmarried partner component are less problematic than the spousal estimates.

This analysis indicates the difficulties associated with counting the number of same-sex spouses as long as data capture errors unavoidably constitute a large component of the group to be

³⁴ See page 23 in the following report for the definition of a PUMA and how to use this geographical concept. <http://www.census.gov/acs/www/Downloads/handbooks/ACSstateLocal.pdf>

DRAFT

identified. While attempts should be made to improve the collection of data for same-sex couple households, an equally important part of any future program must address the issue of reducing data capture errors for responses made to the gender item for opposite-sex couple households that inadvertently contribute to the overall numbers of same-sex couple households. Improvements in upcoming surveys—such as the use of computerized instruments with internal checking routines--will minimize errors caused by paper and pencil formats and will produce beneficial impacts on the quality of data being collected.

APPENDIX A. RELATIONSHIP ERRORS IN TABULATING SAME-SEX COUPLE HOUSEHOLDS

While this paper has focused on data capture errors in the gender item, mention should be made of possible errors in the other item used to identify same-sex couples, namely the relationship item. In both the 2000 and 2010 Census and the ACS, the first category in the relationship item was labeled “Husband/wife” and the second category beneath it was “Biological son or daughter.” If there were significant problems created by incorrectly marking the husband/wife box on the form instead of the son or daughter box, then this would be revealed in the age differences of the couples. The householder would be 20-30 years older on average than the partner, if that same-sex partner was really a son or daughter.

DRAFT

Using Texas as an example, Appendix Table A clearly shows that this is not the case. For all categories of same-sex couples, the householder is usually 1 to 2 years older than the partner as is typical among opposite-sex couples. This would not be found if large proportions of same-sex couples really represented householders and their children. In fact, if one examines the same-sex couple data lines for those with name-sex inconsistencies, the average ages of these couples clearly is closer to the population that they probably came from due to data capture errors, namely opposite-sex couples.

Appendix Table A. Average Ages of Householders and Partners for Couples: Texas, 2010 Census

| Type of living arrangement | Age of householder (in years) | Age of partner (in years) | Difference in ages householder-partner (in years) |
|-----------------------------------|-------------------------------|---------------------------|---|
| Same-sex unmarried partner | 42.6 | 40.4 | 2.2 |
| No name-sex inconsistency | 42.9 | 40.7 | 2.2 |
| Name-sex inconsistency | 38.5 | 37.4 | 1.1 |
| Opposite-sex unmarried partner | 36.9 | 35.7 | 1.2 |
| Same-sex spouse | 48.0 | 46.5 | 1.5 |
| No name-sex inconsistency | 46.0 | 44.5 | 1.5 |
| Name-sex inconsistency | 48.9 | 47.4 | 1.5 |
| Opposite-sex spouse | 49.0 | 47.2 | 1.8 |

DRAFT

APPENDIX B. FEMALE-FEMALE FIRST NAME INDICES FOR TEXAS FROM THE 2010 CENSUS

Appendix Table B. First Name Indices for Female-Female Households in Texas, 2010 Census

(Numbers of reported partners)

Partner Name Index

| HHldr Index | 0-50 | 51-100 | 101-499 | 500-899 | 900-949 | 950-1000 | NA | Total |
|-------------|--------|--------|---------|---------|---------|----------|-------|--------|
| 0-50 | 14,690 | 349 | 1,010 | 460 | 164 | 3,021 | 1,217 | 20,911 |
| 51-100 | 335 | 14 | 16 | 11 | 9 | 65 | 32 | 482 |
| 101-499 | 1,092 | 44 | 113 | 43 | 14 | 172 | 109 | 1,587 |
| 500-899 | 606 | 24 | 67 | 56 | 8 | 56 | 83 | 900 |
| 900-949 | 279 | 5 | 25 | 9 | 1 | 3 | 39 | 360 |
| 950-1000 | 6,982 | 152 | 358 | 108 | 13 | 53 | 466 | 8,132 |
| NA | 1,121 | 26 | 93 | 52 | 17 | 212 | 1,757 | 3,278 |
| Total | 25,104 | 614 | 1,682 | 739 | 226 | 3,582 | 3,703 | 35,650 |

Red lined rows/columns indicate that either the householder or partner has a name likely to be a male first name. Green cell (14,690) indicates 41 percent of partners both have names that are female at least 95 percent of the time (having a male name index of 0-50).

DRAFT

APPENDIX FIGURE 1. MAILOUT/MAILBACK 2010 CENSUS FORM D-1 (VERTICAL OR SEQUENTIAL
FORMAT DESIGN)

DRAFT

DRAFT

REFERENCES

Kathleen T. Ashenfelter, "Eye-tracking Study Report: Examining User Patterns for Demographic Items on the 2007 and 2008 Mail Forms" Survey Methodology Study Series (SSM 2010-01), Statistical Research Division, U.S. Census Bureau, January, 2010
<http://www.census.gov/srd/papers/pdf/ssm2010-01.pdf>

Gary Gates, "Same-sex couples in US Census Bureau Data: Who gets counted and why." The Williams Institute, UCLA School of Law, August 2010
http://services.law.ucla.edu/williamsinstitute/pdf/WhoGetsCounted_FORMATTED1.pdf

Gary Gates and Michael D. Steinberger, "Same-Sex Unmarried Partner Couples in the American Community Survey: The Role of Misreporting, Miscoding and Misallocation." Paper presented at the Annual Meetings of the Population Association of America, Detroit, MI, April 30, 2009.

Martin O'Connell and Gretchen Gooding, "Editing Unmarried Couples in Census Bureau Data." July 2007, United States Census Bureau working paper.
<http://www.census.gov/population/www/documentation/twps07/twps07.html>

Martin O'Connell, Daphne Lofquist, Tavia Simmons, and Terry Lugaila, "New Estimates of Same-Sex Couple Households from the American Community Survey." Paper presented at the Annual Meetings of the Population Association of America, Dallas, Texas, April 15-17, 2010.
http://www.census.gov/population/www/socdemo/hh-fam/SS_new-estimates.pdf

Paula J. Schneider, *Content and Data Quality in Census 2000*, Census 2000 Testing, Experimentation, and Evaluation Program Topic Report No. 12, TR-12 (US Census Bureau: Washington DC, 2004), Table 1.

Williams Institute, "Williams Institute Experts Comment on Department of Justice DOMA Decision." February 24, 2011 <http://www3.law.ucla.edu/williamsinstitute/pdf/Pressrelease2.24.pdf>