



Sample Errors Call Into Question Conclusions Regarding Same-Sex Married Parents: A Comment on “Family Structure and Child Health: Does the Sex Composition of Parents Matter?”

D. Paul Sullins¹

Published online: 2 October 2017

© The Author(s) 2017. This article is an open access publication

Abstract Because of classification errors reported by the National Center for Health Statistics, an estimated 42 % of the same-sex married partners in the sample for this study are misclassified different-sex married partners, thus calling into question findings regarding same-sex married parents. Including biological parentage as a control variable suppresses same-sex/different-sex differences, thus obscuring the data error. Parentage is not appropriate as a control because it correlates nearly perfectly ($+0.97$, gamma) with the same-sex/different-sex distinction and is invariant for the category of joint biological parents.

Keywords Family structure · Marriage · Same-sex families · Child health · Data integrity

Introduction

In the study “Family Structure and Child Health: Does the Sex Composition of Parents Matter?,” Reczek and colleagues presented a well-constructed analysis of health measures from the National Health Interview Survey (NHIS) for children with parents in four union statuses: same-sex married, same-sex cohabiting, different-sex married, and different-sex cohabiting (Reczek et al. 2016). However, the National Center for Health Statistics (NCHS), which administers the NHIS, has advised that for a substantial proportion of the data included in the sample, a large majority of the cases classed as same-sex married partners are believed to be misclassified different-sex married

The original article can be found at <http://dx.doi.org/10.1007/s13524-016-0501-y>

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13524-017-0616-9>) contains supplementary material, which is available to authorized users.

✉ D. Paul Sullins
sullins@cua.edu

¹ Department of Sociology, The Catholic University of America, Washington, DC 20064, USA

partners (National Center for Health Statistics 2015). Including these cases resulted in the probable underestimation of differences between different-sex and same-sex married partners in this study and brings any related conclusions into question. The problem affects only same-sex married partners, not same-sex cohabiting partners, and only outcomes for which their children may differ from those of different-sex married partners in the population. Herein, I illustrate and discuss the effect of the corrupted sample on findings for child emotional difficulties, one of five outcomes examined in the original article.

NCHS reported the classification problem in April 2015, advising that “a change in data verification procedures from 2004 to 2007 may have led to uncorrected errors in identification . . . [such that] some opposite-sex married couples are believed to have been incorrectly identified as same-sex for data years 2004–2007” (National Center for Health Statistics 2015:2). Moreover, “the number of incorrect sex codes identified through [a subsequent verification process] is large compared to the number of married couples initially identified as being same-sex” (p. 6). After discovering that errors may exist, an NCHS review of the problem cases’ sex designations based on first names concluded that more than three-quarters of the cases were incorrectly identified; further review estimated that from 66 % to 84 % of initially identified same-sex married couples were erroneously classified different-sex married couples. The NCHS report advises that because “these errors will affect analyses of this subgroup for the relevant years,” NCHS “discourages analyses of this subgroup for those data years” (that is, 2004–2007, quarter 2; hereafter referred to as the “embargoed period”) (National Center for Health Statistics 2015:6).

Reczek et al. (2016) reported that their same-sex married parents analysis sample consisted of 216 cases from 2004–2013 combined data, which therefore included approximately 117 cases from the embargoed period. This is evident from the numbers reported as follows (see Table 1 for reference): NCHS reports identifying, by spousal and sex matching, a total of 342 same-sex married couples from 2004–2011, of which only 66 were reported outside the NCHS embargoed period. My own independent identification results in totals very similar to NCHS totals for these years, and totals identical to theirs for all years subsequent to the embargoed period for which they provide comparison numbers. Extending the NCHS total to 2013 by my count yields a total of 412 cases of same-sex married couples from 2004 to 2013, of which only 136 were outside the embargoed period. Note that this is the total of same-sex married *couples*—not just parents, which is a much smaller number. Because Reczek et al. reported analyzing more same-sex married parents (216) than there are same-sex married couples (136) outside the embargoed period, they must have included the cases from the embargoed period. My estimate of the number of affected cases is derived by interpolation from the distribution of same-sex married partners by year to the distribution of same-sex married parents. The actual number of misclassifications is unknown and may be more than or less than this estimate. If the error rate (77 %) identified in the NCHS names review is accurate, then approximately 90 cases, or 42 % of Reczek et al.’s analysis sample of “same-sex married parents,” consists of misclassified different-sex married parents. Moreover, NCHS’s count and mine agree in identifying only 148 cases of same-sex married parents from 2004–2013 (including the embargoed period), while Reczek et al. claimed 216 for this period, suggesting potential further misidentifications. Although Reczek et al. discussed measurement

Table 1 Same-sex married (spousal partner) cases, by year: NHIS 1997–2015

	Same-Sex Married Couples		
	Total per Sullins (NCHS)	With Children	With Sample Child Data
1997–2003	39 (49)	7	6
Embargoed Period			
2004	112 (111)	48	44
2005	64 (68)	28	22
2006	67 (68)	31	27
2007	29 (29)	14	14
2008	10 (10)	2	2
2009	14 (14)	1	1
2010	24 (24)	6	6
2011	18 (18)	1	1
2012	26	7	6
2013	44	10	10
2014	80	18	17
2015	111	24	23
Total	603	190	173
Total 2008–2015 (non-embargoed)	327	69	66

issues briefly in the article, they did not mention the embargo or explain the disparity between their count and that of NCHS, which probably was not published until after the article had already been submitted for publication.

For any comparison between two groups A and B, the effect of erroneously classifying a substantial proportion of cases from group A as belonging to group B will necessarily reduce the observed magnitude of any true differences between A and B, with corresponding effects on prevalence rates, ratios, and related inferences. As Black et al. (2007) explained in greater detail with regard to a similar problem on the 2000 census, because same-sex married couples account for only a fraction of 1 % of all married couples in the U.S. population during the period studied, even a small random error in sex designation can result in a large inaccuracy in specifying the members of this tiny subpopulation, resulting in consequent false positives or negatives on hypothesis tests.

Tables 2 and 3 illustrate the probable effect of including the embargoed cases on Reczek et al.'s reported findings. The outcome variable in the tables is emotional difficulties, which is used only on the in-depth sample (sample child file). The replication model in Table 2 presents unadjusted odds ratios that replicate, as far as possible, Reczek et al.'s data and analysis method and are very similar to their baseline findings reported in Model 1 of Table 5. A model that imposed demographic and economic controls similar to theirs (not shown) also produced findings similar to those of Model 2 in Table 5 in their study, which also imposes economic controls, but substantially reduced the effective n because of missing data. The replication model coefficients generally support the hypothesis conclusions drawn in the article to the same extent as the baseline model: that is, children's health is generally similar among

Table 2 Odds ratios for emotional difficulties by family structure, replicating Reczek et al.: NHIS 2004–2013 ($n = 89,435$)

	Baseline: Reczek et al. Table 5, Model 1	Replication
Family Structure (ref. = man-woman married; $n = 82,585$)		
Same-sex married ($n = 133$)	1.02	0.98
Same-sex cohabiting ($n = 265$)	3.04 ^{a***}	3.58 ^{a***}
Different-sex cohabiting ($n = 6,382$)	1.82 ^{***}	1.74 ^{***}
Support for H1	Yes	Partial
Support for H2	Yes	Yes

Notes: Models report survey-weighted odds ratios from ordered logistic regressions. Asterisks report significance of t test for difference from reference group.

^a Different from same-sex married ($p < .05$).

*** $p < .001$

children whose parents have the same union status, regardless of whether those parents are same- or different-sex (Hypothesis 1 (H1)), and parental marriage is associated with healthier outcomes for both same- and different-sex parents (Hypothesis 2 (H2)). In the replication model, as in the baseline model, same-sex married outcomes do not differ significantly from different-sex married (supporting H1) but do differ significantly from same-sex cohabiting (supporting H2). Support for H1 is qualified in the replication model by the fact that same-sex cohabiting differs significantly from different-sex cohabiting, which is the case for the baseline model after the inclusion of economic controls (Model 2 of Table 5) in the article. I report this outcome for the replication model as partial support for H1.

Table 3 illustrates the effect of including and excluding the embargoed period on model outcomes and related hypothesis assessments. H1 and H2 are cited in full in the table notes for convenience. The first column in panel 1 shows the same replication model already presented in Table 2, using the same data as did Reczek et al.: that is, 2004–2013 cases including the embargoed period. Panel 2, which differs only by excluding the embargoed period, finds relative odds for emotional difficulties with same-sex married parents that are twice as high as those in panel 1, while estimates for cohabiting partners are unaffected. These effects are consistent with the problems identified by NCHS, which affected only the classification of same-sex married partners. In panel 2, same-sex cohabiting no longer differs statistically from same-sex married, thus not supporting H2, although same-sex married still does not differ statistically from different-sex married. However, with only 26 cases, it may be that the same-sex married sample lacks sufficient power to discriminate this difference if it existed.

Panel 3 examines this possibility by extending the sample capture by two more years so as to include additional cases. This increases the size of the same-sex married sample to 66; the additional cases have all been verified using the procedure instituted after 2007. In panel 3, the replication model does not support H1 or H2. Relative odds for emotional problems with same-sex married and same-sex cohabiting parents do not differ from each other, and are both significantly higher than with their different-sex counterparts.

Table 3 Odds ratios for emotional difficulties by family structure, illustrating the effect of the embargoed period: NHIS 2004–2013 (panels 1 and 2) and 2004–2015 (panel 3)

Replication Model Using:	Emotional Difficulties		Including Biological Relatedness	
	Per Reczek et al.	Per CDC Usage	Per Reczek et al.	Per CDC Usage
1. 2004–2013 Cases Including the Embargoed Period				
Family structure (ref = man-woman married; $n = 82,585$)				
Same-sex married ($n = 133$)	0.98	1.97	0.55	0.92
Same-sex cohabiting ($n = 265$)	3.58 ^{a,b***}	4.21 ^{a***}	2.08 ^{b**}	2.12 ^{**}
Different-sex cohabiting ($n = 6,382$)	1.74 ^{***}	1.77 ^{***}	1.39 ^{***}	1.34 ^{**}
Support for H1	Partial	Partial	Yes	Yes
Support for H2	Yes	Different-sex only	Yes	Different-sex only
2. 2004–2013 Cases Excluding the Embargoed Period				
Family structure (ref. = man-woman married; $n = 82,585$)				
Same-sex married ($n = 26$)	2.13	6.29 ^{**}	1.26	2.30
Same-sex cohabiting ($n = 265$)	3.58 ^{a***}	4.21 ^{a***}	2.88 ^{a***}	2.12 ^{**}
Different-sex cohabiting ($n = 6,382$)	1.74 ^{***}	1.77 ^{***}	1.78 ^{***}	1.34 ^{**}
Support for H1	Partial	No	Partial	Yes
Support for H2	Different-sex only	Different-sex only	Different-sex only	Different-sex only
3. 2004–2015 Cases Excluding the Embargoed Period				
Family Structure (ref. = man-woman married; $n = 101,074$)				
Same-sex married ($n = 66$)	2.86 [*]	5.58 ^{***}	1.53	2.60 [*]
Same-sex cohabiting ($n = 331$)	2.95 ^{a***}	3.69 ^{a***}	1.77 ^{**}	1.87 ^{**}
Different-sex cohabiting ($n = 8,081$)	1.69 ^{***}	1.67 ^{***}	1.33 ^{**}	1.23 ^{**}
Support for H1	No	No	Yes	Partial
Support for H2	Different-sex only	Different-sex only	Different-sex only	Different-sex only

Notes: Models report survey-weighted odds ratios from logistic regression models (ordered for “Per Reczek et al.” and binary for “Per CDC Usage”). First column reports unweighted N . H1 hypothesizes that children’s health will generally be similar among children whose parents have the same union status, regardless of whether those parents are same- or different-sex. H2 hypothesizes that for both same-sex and different-sex couples, parental marriage will be associated with healthier outcomes. Superscript letters and asterisks report significance of t test for difference from the reference group.

^a Different from different-sex cohabiting ($p < .05$).

^b Different from same-sex married ($p < .05$).

* $p < .05$; ** $p < .01$; *** $p < .001$

Columns 2 and 4 of Table 3 present for comparison a more common and possibly more useful treatment of the variable that Reczek et al. labeled “emotional difficulties.” Response options for the corresponding NHIS question, “Overall, do you think that [sample child] has any difficulties in one or more of the following areas: emotions, concentration, behavior, or being able to get along with other people?” were (1) “no”;

(2) “yes, minor difficulties”; (3) “yes, definite difficulties”; or (4) “yes, severe difficulties.” Rather than fitting a proportional odds ratio to all four categories (three steps), as did Reczek et al., the Centers for Disease Control and Prevention (CDC) morbidity and prevalence reports combined the top two categories (definite and severe) into “serious” difficulties, which is contrasted with the combined bottom two categories (none or minor) using binary logistic regression models (Blackwell 2010; Pastor et al. 2012). This classification was adopted following a detailed evaluation of the measure’s psychometric properties compared with other evaluative instruments and use of mental health services (Bourdon et al. 2005), which concluded “that the measure of serious overall difficulties is an effective and efficient screener for child and adolescent mental health problems in the United States” (Pastor et al. 2012:3). In other words, for assessing child mental health problems in the population from parent reports, including the distinctions between no difficulties and minor difficulties, and between definite and severe difficulties, is less dispositive than distinguishing between none or minor difficulties on the one hand and definite or severe difficulties on the other hand.

The Table 3 columns headed “Per CDC usage” present this alternative method and classification. The odds ratios in these columns are susceptible to a more direct interpretation in terms of relative risk, and are directly comparable with published reports of child mental health problems prevalence or risk in other U.S. subpopulations (Blackwell 2010; Perou et al. 2013). For the purposes of this comment, they also appear to be somewhat more sensitive to distinctions among the different sample configurations and lead to more determinative conclusions regarding hypothesis assessment. Using the CDC categories, the estimated risk of child emotional problems with same-sex married parents in the replication model (second column of Table 3) increases by a factor of 3 with the exclusion of the embargoed period (from panel 1 to panel 2), more than in the first column. This model in panel 2 also already substantiates, without additional cases, the hypothesis assessments reached in panel 3 using proportional odds. With the corrected data, the CDC method risk factors result in point estimates of serious child emotional difficulties that are lower for same-sex cohabiting than for same-sex married—an outcome quite different than found in the original article. This difference is not statistically significant (although it is less than .10 in column 2, panels 2 and 3), resulting in no change in hypothesis evaluation, but it may alter the theoretical consideration of H2 to observe that child outcomes with same-sex parents are not significantly worse if their parents are married than to observe that their outcomes are not significantly better. Comparing panels 2 and 3, the CDC method risk factors also suggest that the risk of emotional problems with same-sex parents may be declining, perhaps because of increasing social or legal acceptance. Nonetheless, the overall results for the CDC metric lead to the same hypothesis evaluations as does Reczek et al.’s method for the larger corrected sample (panel 3), confirming the general accuracy of their analysis, but they reverse the original conclusions of the article regarding H1 and H2 for this outcome.

Reczek et al. advised that they attempted to certify their sample by examining different sample configurations. Their ability to observe the differences among these samples may have been impeded by their inclusion of a “family relatedness” control variable derived from an NHIS measure of biological parentage—that is, whether the child being observed resided with both biological parents, only one biological parent, or neither biological parent. Inclusion of this control moderates the odds on emotional difficulties for both same-sex union statuses much more than for different-sex partners.

As Table 3 shows (comparing columns 1 and 3 or columns 2 and 4), relative to the comparable unadjusted model, including biological relatedness reduces odds for same-sex couples by roughly one-half (40 % to 63 %), but for different-sex couples by at most one-quarter (–2 % to 24 %). After biological relatedness is controlled for, none of the odds for same-sex married relative to different-sex married are significant using Reczek et al.’s classification, thus supporting H1. But this finding reflects a misuse of this control variable.

Biological parentage cannot serve as a control variable when comparing same-sex and different-sex parents because (as the article acknowledges) same-sex partners cannot be joint biological parents. For same-sex partners, the category “two biological parents” is not a variable but rather a constant because it cannot be other than null. Table 4 shows the proportions for all categories of parentage by union status on NHIS. For two groups being compared, a condition that is not possible for one group—and in this case is the modal condition for the other group—cannot serve as a control; rather, such a condition is part of the definition, or at least highly correlated with the defined difference, between the two groups. Controlling for biological parentage when analyzing differences due to same-sex or different-sex parent partnership is like controlling for skin color when analyzing black-white race differences, or the presence of a Y chromosome when analyzing sex differences. The “control” in these cases acts as a suppressor variable masking most or all differences in the variable of interest. It is not that these variables are unimportant; but they are really *explanatory* variables, not *control* variables, as must also be the case for biological parentage with regard to same-sex/different-sex analyses.

Put another way, biological parentage violates the assumption of independence for regression predictors because it is highly correlated with any differences that may be due to having same-sex versus different-sex parents. The correlation (gamma) between the three categories of biological parentage and same-sex/different-sex parents on NHIS, shown in Table 4, is a nearly perfect inverse .97 for married parents and a positive .98 for cohabiting parents. Reczek et al. combined two- and one-biological-parent categories and included an “unspecified parentage” category, perhaps in an attempt to overcome this problem. The resulting three categories have an overall correlation with same-sex/different-sex parents near 0, but this is only a resultant of the fact that the first two

Table 4 Weighted proportions and means for family biological relatedness by union status: NHIS 1997–2012 ($n = 89,435$)

	Same-Sex		Opposite-Sex	
	Married ($n = 72$)	Cohabiting ($n = 439$)	Married ($n = 153,313$)	Cohabiting ($n = 17,960$)
Gamma (same-sex)			–0.97	0.98
Family Relatedness (biological parentage)				
Two biological parents	0***	0***	85.4	68.4***
One biological parent	76.2***	79.4***	13.2	30.9***
No biological parents	23.8**	20.6***	1.3	0.7***

Notes: Same-sex married cases embargoed by CDC because of probable assignment errors (2004–2007 second quarter) are excluded. Asterisks indicate significance of t test for difference from the reference group.

** $p < .01$; *** $p < .001$

categories have a nearly perfect positive correlation (gamma .98), and “unspecified parentage” simply reflects the proportion of cases in 2013, when NCHS changed the wording of the question, which has nothing to do with a substantive difference in parentage. Regardless of whether one agrees that biological parentage is theoretically relevant, the pertinent point here is that attempting to equalize or control for biological parentage obscures a highly influential class of difference between the two groups.

It is no criticism of an otherwise careful analysis to acknowledge that, through no fault of the authors, the data contained errors of which they were unaware. My own study of same-sex partners using NHIS also inadvertently used data from the embargoed period, leading to the underestimation of emotional problems risk (Sullins 2015a); a correcting note is published at Social Science Research Network (SSRN) (Sullins 2015b). Support for my conclusions, which hypothesized difference, was thereby strengthened, but underestimation would likely have the opposite effect for Reczek et al.’s hypotheses positing similar outcomes for children with same-sex and different-sex married parents (H1 and H2). This comment has pertained to only one of five outcomes, and two of four hypotheses, examined in their study. Although using a sample that excludes the embargoed period cases will very likely reduce same-sex married odds ratios for the other outcomes, this change may or may not affect assessments of their hypotheses. Perhaps, using the identifications presented in the next paragraph or some other, they would be willing to show the effect, if any, of using the corrected sample on the remaining outcomes and hypotheses examined in their study.

Researchers addressing questions regarding same-sex married couples using partner matching on NHIS should avoid including cases from the embargoed period, following the NCHS’s disclosure of the high likelihood of misspecification error that may lead to false conclusions. To facilitate more precise and general use of these data, with this comment I am publishing NHIS same-sex partner identification variables for 1997–2015, with flags for excluding the embargoed period for spousal partners, derived from my own analysis. As reported in Table 1, the annual totals for these variables closely match those published by NCHS for all NHIS annual data since 1997 and are identical to them since 2008. However, NCHS has no responsibility for these identifications or any errors therein. It is hoped that these identification variables will encourage more accurate investigation, from any theoretical perspective, into the characteristics and concerns of same-sex married couples and their children.

Acknowledgments Partial funding for the research underlying this comment was provided by Research Grant 470225 from the Catholic University of America.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Black, D., Gates, G., Sanders, S., & Taylor, L. (2007). *The measurement of same-sex unmarried partner couples in the 2000 U.S. Census* (Working Paper Series No. CCPR-023-07). Los Angeles: California Center for Population Research. Retrieved from <https://escholarship.org/uc/item/72r1q94b.pdf>

- Blackwell, D. (2010). *Family structure and children's health in the United States: Findings from the National Health Interview Survey, 2001–2007* (Vital Health Statistics Series 10, No. 246). Hyattsville, MD: National Center for Health Statistics.
- Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44, 557–564.
- National Center for Health Statistics. (2015). *Changes to data editing procedures and the impact on identifying same-sex married couples: 2004–2007 National Health Interview Survey* (National Health Interview Survey April 2015 Report). Hyattsville, MD: National Center for Health Statistics. Retrieved from <https://www.cdc.gov/nchs/data/nhis/samesexedits.pdf>
- Pastor, P. N., Reuben, C. A., & Duran, C. R. (2012). *Identifying emotional and behavioral problems in children aged 4–17 years: United States, 2001–2007* (National Health Statistics Reports No. 48). Hyattsville, MD: National Center for Health Statistics.
- Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., . . . Huang, L. N. (2013). *Mental health surveillance among children—United States, 2005–2011* (Morbidity and Mortality Weekly Report, Vol. 62, No. 2 Suppl.). Washington, DC: Centers for Disease Control and Prevention.
- Reczek, C., Spiker, R., Liu, H., & Crosnoe, R. (2016). Family structure and child health: Does the sex composition of parents matter? *Demography*, 53, 1605–1630.
- Sullins, D. P. (2015a). Emotional problems among children with same-sex parents: Difference by definition. *British Journal of Education, Society & Behavioural Science*, 7, 99–120.
- Sullins, D. P. (2015b). *Emotional problems among children with same-sex parents: Difference by definition* (SSRN Scholarly Paper No. ID 2500537). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=2500537>