TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 64

Supplementary Materials for "Same-Gender Sex among U.S. Adults: Trends across the 20th century and during the 1990s"

Charles F. Turner, Maria A. Villarroel, James R. Chromy,
Elizabeth Eggleston, Susan M. Rogers

Reference Citation

Turner CF, Villarroel MA, Chromy JR, Eggleston E, Rogers SM. (2004) Supplementary Materials for "Same-Gender Sex among U.S. Adults: Trends across the 20th century and during the 1990s" . *Technical Papers on Health and Behavior Measurement, No. 64.* Washington DC: Program on Health and Behavior Measurement, Research Triangle Institute.

Program in Health and Behavior Measurement
Research Triangle Institute

E-Mail
Measurement@RTI.org

SUPPORTING MATERIAL

FOR

Same-Gender Sex among U.S. Adults

Trends across the 20th century and during the 1990s

Charles F. Turner, Maria A. Villarroel, James R. Chromy,

Elizabeth Eggleston, Susan M. Rogers

Supplementary Information on

Data and Methods

Sample Design and Execution. Data were collected in 10 rounds of the General Social Survey (GSS)[1] conducted between 1988 and 2002 (N = 9,487 males, and 12,336 females). (GSS surveys were conducted annually between 1988 and 1994 with the exception of 1992 and biennially since 1996.) The GSS conducted in-person surveys of multistage area probability samples of U.S. households and randomly selected one English-speaking adult (aged 18 and older) for interview at each household that was successfully screened. GSS interviews last approximately 90 minutes each and cover a wide range of topics of interest to social and behavioral scientists. During the period 1988 to 2002 the GSS obtained a median survey response rate of 76.7% (range: 70.0% to 82.4%).[2]

Measurements of Same-Gender Sexual Contact. All ten GSS surveys[3] conducted between 1988 and 2002 included questions that allow us to identify respondents who reported having a same-gender sexual partner in the past 12 months and, from the 1989 survey onwards, a same-gender sexual partner since age 18. From 1991 onwards, the GSS also included questions

that permit us to identify respondents who reported having a same-gender sex partner in the past five years. Question wordings were:

> *Since Age 18:* "Now thinking about the time since your 18th birthday (including the past 12 months) how many female [male] partners have you had sex with?" (Respondent supplied a number.)
>
> *For Past five years:* "Now think about the past five years--the time since [Month/Year], and including the past 12 months, how many sex partners have you had in that five year period? No partners, 1 partner, 2 partners, 3 partners, 4 partners, 5 - 10 partners, 11 - 20 partners, 21 - 100, partners, more than 100 partners?" Have your sex partners in the last five years been exclusively male, both male and female, exclusively female?"
>
> *For Past Year:* "How many sex partners have you had in the last 12 months? No partners, 1 partner, 2 partners, 3 partners, 4 partners, 5 - 10 partners, 11 - 20 partners, 21 - 100, partners, more than 100 partners?" " Have your sex partners in the last 12 months been exclusively male, both male and female, exclusively female?"

Other Measurements. Our analyses use a range of sociodemographic variables to define subpopulations for which prevalences will be estimated and to impute missing data. We also use answers to two questions on attitudes toward same-gender sex to assist in imputation and to understand trends in reported prevalence over 20th century birth cohorts. Those questions are:

> What about sexual relations between two adults of the same sex — do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?

And what about a man who admits that he is a homosexual ... Should such a person be allowed to teach in a college or university, or not?

Supplementary Results

Impact of Missing Measurements. 3,252 of the 21,823 GSS respondents (14.9%) did not answer the questions required to determine if they had a same-gender sex partner during the 12 months prior to the survey. Similarly high rates of missing data afflicted measurements of same-gender contacts in the past five years (15.9%) and since age 18 (21.4%).[4] While a relatively high incidence of missing data is often reported in surveys using paper-and-pencil self-administered questionnaires (SAQs),[5] it does raise concerns about potential nonresponse bias in the GSS measurements. This concern is heightened by the fact that the number of GSS respondents who did not respond to these questions substantially exceeded the number who reported same-gender sexual contact in response to the questions.

To examine the potential for bias introduced by this nonresponse, we used the available data to estimate the odds that missing cases would have reported same-gender contacts if they had completed the survey. Logistic regression models were estimated separately for males and females (using the variables shown in Table 3 as predictors). The first model used all variables except number of heterosexual partners[6], and attitudes toward homosexual sex and allowing gay men to teach in college. A second model included all predictor variables except number of heterosexual partners. All variables were treated as categorical in order to capture non-monotonic associations between predictor and dependent variables.

This estimation procedure was used to impute the probability of reporting a same-gender sexual partner since age 18, in the past year, and in the past five years for cases in which no data were reported. For cases with complete data on the predictor variables and missing data on

same-gender contact, the estimated probability of contact was substituted for the missing observation, and the resulting prevalence estimate was compared to that obtained from the reported data alone. This imputation procedure yielded prevalence estimates that were very similar to those obtained using the reported data alone.[7] In no case did addition of the imputations change the estimated prevalence by more than 0.1 percentage points.

A multiple imputation algorithm developed by Raghunathan et al.[8] et al. was also used both to impute missing values and to obtain an estimate of sampling error that incorporated the effect of the added error due to imputation.[9] The prevalence estimates obtained using this algorithm were also highly similar to those obtained using the reported data alone. No deviation was greater than 0.25 percentage points.

The substantial similarity of these alternative estimates suggests that non-response to these questions can be classified as substantively ignorable — at least to the extent captured by our models. Given the similarity of prevalence estimates derived with and without those imputations, we confine subsequent analyses to respondents' reports of same-gender sex without imputation of missing responses.

Consistency of Sexual Behavior Measurements. Measures of the consistency of respondents' reports of same-gender sex are calculated from three questions. Respondents in all GSS surveys between 1988 and 2002 were asked to report the number of sex partners they had in the preceding twelve months and the gender of those partners (all male, all female, or both). Beginning in 1991, respondents were asked a parallel pair of questions about their sex partners during the past five years. Beginning in 1989, respondents were also asked to report separately the number of male and female sex partners they had since age 18. While these questions obviously do not measure the same characteristic, they offer opportunities for logically

inconsistent reporting (e.g., reporting no same-gender partners in the past five years but both male and female partners during the past year).

Supplemental Table A shows the number of comparisons that could be made for the three possible pairings of measurements (since age 18 vs. past five years; since age 18 vs. past year; past five years vs. past year) and the percent of responses that were logically consistent. All three measurements were coded as binary variables indicating whether the respondent reported any same-gender sex during the time period. A pair of responses was considered logically consistent if the respondent reported: (1) no same-gender sex during either time period, (2) same-gender sex during both time periods, or (3) same-gender sex during the longer time period (e.g., since age 18) but not during the more recent time period (e.g., in past year). Responses were considered logically inconsistent if the respondent reported same-gender sex during the more recent period (e.g., past year) but not during the longer time period (e.g., in past five years or since age 18). If a measurement was missing for one (or both) of the time periods in a comparison, that comparison was excluded from the analysis. Comparison of reports of same-gender sex since age 18 vs. in the past five years excluded persons ages 18 to 22 since their "past five years" would include time periods before their 18th birthday. For the same reason, 18 year olds were excluded from the comparison of reports for the past year vs. since age 18.

Supplemental Table A indicates that across all respondents, there was an extremely high level of consistent reporting of same-gender sexual contact. More than 99% of comparisons were logically consistent; however, of 41,542 pairwise comparisons, 95% reported no same-gender sexual contact in either time period being compared.)

The second panel of Supplemental Table A restricts our analysis to persons who reported same-gender sex in the more recent of the two time periods (i.e., *past year* when compared to past five years; and *past year and past five years* when compared to since age 18). In these

instances, respondents must also report same-gender sex in the longer time period to be logically consistent. Using this more stringent test, we still find that 92.7% of all reports of same-gender sex are logically consistent, but this analysis also reveals that women were slightly more consistent than men in their reporting (95.1% vs. 90.5% unweighted; $p < 0.01$).

Consistency of Gender Measurements. Black et al.[10] have noted that errors in recording respondents' gender could have a substantial impact on estimates of the prevalence of same-gender sex. The GSS includes two indicators of respondent gender. The first is the interviewer's observation of the respondent's gender.[11] In addition, interviewers constructed a household roster that included the age, name, relationship to head of household, marital status (if age 13 or older), and gender of every household member. Interviewers subsequently recorded the line number in the household roster for the respondent. Across the 1988 to 2002 surveys, there were disagreements between these two indicators for 806 of 21,668 respondents.[12]

Smith[13] examined 20 instances in the 1991 GSS in which these two indicators were discrepant and paper copies of the interview materials were available for review. In 18 of 20 cases the interviewers' observations were judged to be accurate. (In 13 cases, the discrepancy was due to mis-recording of the roster line number and in 5 cases the wrong gender was recorded on the household roster.) In only 2 of these 20 discrepancies were the interviewers' observations judged to be inaccurate. If this result were true across all years of the 1988-2002 GSS, it would translate into 81 gender errors for the 21,668 GSS respondents (i.e., 10% of the 806 observed discrepancies).

Impact of Measurement Errors. We performed additional analyses to assess the potential impact upon our findings of measurement error in recording of respondent gender or reporting of same-gender sexual contacts. Supplemental Tables B and C presents alternative estimates of the prevalence of same-gender sexual contacts derived when the sample is purged of

cases for which there is either logically inconsistent reporting of same-gender sexual contacts or disagreement between the interviewer's recording of the respondent's gender and the gender derived using the household enumeration form. (Cases were also excluded if gender was not coded on the Household Ennumeration Form or there was not at least one logical consistency check that could be performed on the reporting of same-gender contacts.)[14]

Supplemental Tables B and C demonstrate that purging the database of these cases does not alter our substantive conclusions regarding variation in the prevalence of same-gender sex reported across birth cohorts or during the 1990s. Thus females continue to have a significant and substantial linear trend of increased reporting of adult same gender contacts ($p < 0.0001$) across birth cohorts while men evidence no similar trend ($p = 0.263$). Similarly, reporting of same-gender contacts during adulthood is significantly higher for females in the 1996-2002 surveys than in 1989-1994 (OR = 1.52, $p = .0007$), but this is not true for males (OR = 1.14, $p = 0.298$). Females in the 1996-2002 surveys also reported a higher prevalence of same-gender contacts in the past five years (OR = 1.70, $p = .006$) and past year (OR = 1.96, $p = 0.0007$) compared to reporting in earlier surveys. Men show a similar trend for reporting of same-gender contacts in the past year (OR = 1.63, $p = 0.0076$) but not for contacts in the past five years (OR = 1.23, $p = 0.237$)

References and Notes

1 Davis JA, Smith TW., General Social Surveys, 1972-2002. [machine-readable data file]. Principal Investigator, James A. Davis; Director and Co-Principal Investigator, Tom W. Smith; Co-Principal Investigator, Peter V. Marsden, NORC ed. Chicago: National Opinion Research Center, producer, 2002; Dataset supplied by the Inter-University Consortium for Political and Social Research, University of Michigan.

2 Response rate is defined as "response rate 5 (RR5)" in: American Association for Public Opinion Research (AAPOR), Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Lenexa, Kansas: AAPOR, 2000.

3 Our analysis combines data from two randomly-assigned experimental conditions in the 1988 GSS that tested alternate introductions to the sex questions. (One introduction merely promised confidentiality while the second introduction also linked the questions to the AIDS epidemic.) Smith has reported that "there were no statistically significant effects of the introduction on reports of sexual behavior (p. 316)." In 1991, a randomly assigned subset of GSS respondents was asked to report the number of sex partners they had in the 12 months and the past five years without providing the response categories (0, 1, 2, 3, 4, 5-10, 11-20, 21-100, more than 100 partners) used by the standard question format. Open-ended responses for this subset of 1991 respondents were recoded into the categories used in the standard format question. (See: Smith TW. A methodological analysis of the sexual behavior questions on the General Social Survey. Journal of Official Statistics, 1992, 8:309-325.

4 Reported rates of missing data and subsequent imputations exclude persons surveyed in years when measurement was not made (i.e., 1988 for contact in since 18, and 1988-90 for contact in past five years).

5 Fay RE, et al., op. cit.; Witt MB, Pantula J., Folsom RE, Cox, BG. Item nonresponse in 1988. in Turner CF, Lessler J, Gfroerer J, eds. Survey Measurement of Drug Use: Methodological Issues. (DHSS Publication ADM-92-1929). Washington, D.C.: Government Printing Office, 1992.

6 Imputation models incorporating number of heterosexual partners were discarded as unrealistic since they required the assumption that the patterns of association between predictor variables and the odds of reporting same-gender sex were the same for: (1) persons who were willing to report on both heterosexual and homosexual contacts, and (2) those who reported on heterosexual contacts but did not answer questions on homosexual contact.

7 Model 1 was used to impute 3,118 missing observations for reporting of same-gender sex in the past year (measured in all years, 1988-2002). Estimated prevalences using these imputations (plus the 18,571 cases with responses) were 2.78% for males and 1.88% for females. Estimates derived using only the actual responses were 2.76% for males and 1.90% for females. Highly similar results were obtained for estimates of the prevalence of same-gender sex in the past five years and since age 18 and for the imputations made using Model 2.

Model 1 was also used to impute 4,177 cases with missing responses to the question on same-gender sex partners

since age 18 (asked in 1989 through 2002), and 2,650 cases with missing data for the question on same-gender sex partners in the past five years (asked in 1991 through 2002). Estimated prevalence was extremely similar whether the imputations were included in the calculations or not (in past five years: 3.80% vs. 3.79% for men and 3.04% vs. 3.07% for women; since 18: 5.02% vs. 4.96% for men and 4.59% vs. 4.65% for women. Since the additional predictor variables used in Model 2 had missing measurements due to non-response to the attitude questions and the randomized question inclusion scheme (rotation grouping) used in the GSS, these models could impute fewer cases. We restricted our analysis to those years in which the same-gender sex question was asked and the two (of three) randomly-assigned rotation groups that included the questions on attitudes toward same-gender sex and homosexual college teachers. Using Model 2, the largest absolute difference (0.06 percentage points) between observed and imputed values was found for the measurement of same-gender sex since age 18 among females. 1,265 missing values were imputed for females, and the estimated prevalence calculated from the combination of imputed and reported values (n=7,326) was 4.24%. The estimated prevalence of same-gender sex since age 18 calculated solely from respondents reports (n=5,971) was 4.30%. Overall, while our imputations had an extremely small impact on estimates of the prevalence of same-gender sex, it is noteworthy that in all 6 instances (i.e., 2 prediction models for 3 measurements of same-gender sex) the imputation procedure lowered the estimate of female-female contact and increased the estimate of male-male contact.

8 Raghunathan TE, Solenberger PW, Van Hoewyk J. IVEware: Imputation and Variance Estimation Software. Ann Arbor, Michigan: Survey Methodology Program, Survey Research Center, Institute for Social Research, University of Michigan, 2002.

9 This error was captured by creating 5 imputed data sets and combining the estimate variance among imputations to the average design-based variance estimate within imputation. The algorithm uses an iterative approach to impute missing values for a set of variables where each variable may have some missing values. This allowed for the simultaneous imputation of missing values for all 13 variables shown in Table 3 plus the three measures of same gender sex for all 21,823 survey respondents (for the past year question), for the 17,433 respondents in the 1991-2002 surveys (for the past 5 years question), and for the 20,342 respondents in the 1989-2002 surveys (for the same gender sex since age 18 question). The numbers of cases imputed were 3252, 2767, and 4353 respectively.

10 Black D, Gates G, Sanders S, Taylor L. Demographics of the gay and lesbian population in the United States: Evidence from available systematic data sources. Demography, 37(2):139-154 (2000).

11 The first time the field interviewer was asked to record gender occurred in questioning the household informant who supplied information for the household roster. The interviewer was instructed: “Code Sex [ASK IF NOT OBVIOUS]”. A similar interviewer specification does not appear at the conclusion of the interview when the interviewer codes the gender of the survey respondent. We suspect it is likely that experienced interviewers would carry over the admonition to ask about gender if they were uncertain. However, it is also likely (Tom Smith, personal communication, April 23, 2004) that there would be little room for uncertainty in most cases since the interviewer would have obtained the respondent’s first name, interviewed them for approximately 90 minutes, have the gender variable in the household enumeration to consult if they were uncertain of the respondent’s gender and, for married people, had the first name and gender of the spouse on the household enumeration form. Interviewers would also have had to figure out whether "wife" or "husband" was appropriate in asking married respondents several of the GSS interview questions.

12 In addition to the 886 substantive disagreements, household roster data were missing in 155 cases.

13 Smith TW. A report on the GSS household enumeration variables. GSS Methodological Report No. 73. Chicago: National Opinion Research Center, University of Chicago, 1992.

14 Cases were excluded if there was a logical inconsistency in reporting of gender or sexual contacts or if any of the following conditions were true: (1) gender was missing from the household enumeration form, (2) data were missing for two or more of the three sexual behavior measurements (SGS1Yr, SGS5Yr, and SGS18) , (3) the case was from the 1988 GSS survey when only one SGS1Yr was measured, (4) the respondent was 18 years of age and the only non-missing sexual behavior measurements were SGS18 and SGS1Yr, (5) the respondent was between the ages of 18 and 22, and the only non-missing sexual behavior measurements were SGS18 and SGS5Yr.

Mnemonics for sexual behaviors are: SGS18, Reported having a same-gender sex partner since age 18; SGS1Yr, Reported having a same-gender sex partner in the past 12 months; SGS5Yr, Reported having a same-gender sex partner in the past 5 years.

SUPPLEMENTAL TABLE A. Consistency in reporting of same-gender sexual partnerships since age 18, in past 5 years, and in past year. (Unweighted tabulation from the 1988 - 2002 General Social Surveys dataset)

	All Respondents		Respondents reporting SGS in most recent time period (c)	
SAME-GENDER SEX REPORTS (a)	**N Comparisons**	**% Consistent**	**N Comparisons**	**% Consistent**
Males				
Since 18 vs. Last 5 Years (a)	5,441	99.4%	218	84.4%
Since 18 vs. Last Year (b)	6,779	99.7%	195	90.8%
Last 5 Years vs. Last Year	5,850	99.9%	189	97.4%
All Male Comparisons	18,070	99.7%	602	90.5%
Females				
Since 18 vs. Last 5 Years (a)	7,105	99.8%	217	92.6%
Since 18 vs. Last Year (b)	8,723	99.9%	185	94.6%
Last 5 Years vs. Last Year	7,644	100.0%	168	98.8%
All Female Comparisons	23,472	99.9%	570	95.1%
ALL COMPARISONS				
All Time Periods	41,542	99.8%	1,172	92.7%

NOTE. Cases for which one of the pair of measurements was missing were excluded from comparison. Respondents who reported zero partners (of either sex) in the past 1 or 5 years were not asked the gender of their partners in the past 1 or 5 years. Our analysis treats these respondents as having zero same-gender partners during this time period. (Note that inconsistencies could still occur in reporting if the respondent reported same-gender partners during the past year but no same-gender partners in past 5 years.)

(a) Comparison excludes respondents who were less than 23 years of age at time of interview since their reports could include partners acquired before age 18.

(b) Comparison excludes respondents who were 18 years of age at time of interview since their reports could include partners acquired before age 18.

(c) Comparisons include only persons who reported same-gender sex during the most recent of the two time periods.

SUPPLEMENTAL TABLE B. Estimated prevalence of same-gender sexual contact derived using 1989-2002 GSS database purged of cases with inconsistent reporting of either gender or sexual contact.

	Since 18		Last 5 Years		Last Year	
YEAR	**Female %**	**Male %**	**Female %**	**Male %**	**Female %**	**Male %**
	s.e.	s.e.	s.e.	s.e.	s.e.	s.e.
1994 and earlier (a)	3.4%	4.40%	1.7%	2.6%	1.1%	1.8%
	0.32%	0.40%	0.28%	0.36%	0.18%	0.25%
1996-2002	5.1%	5.00%	2.9%	3.2%	2.1%	2.8%
	0.37%	0.39%	0.27%	0.35%	0.23%	0.32%
OR	**1.52**	**1.14**	**1.70**	**1.23**	**1.96**	**1.63**
Categorical, p	***0.0007***	***0.2980***	***0.0055***	***0.237***	***0.0007***	***0.0076***

NOTE. Cases were not included in the database if responses were not available for at least two sexual behavior questions for which there was a logically required consistency of response or if a gender code was not available from the household ennumeration form. See text note 14.

(a) Sexual contact in the past year was measured in all surveys. Contact since age 18 is available from 1989 onwards, and contact in past five years is available for 1991-2002.

SUPPLEMENTAL TABLE C. Estimated prevalence of same-gender sexual contact since age 18 derived using 1989-2002 GSS database purged of cases with inconsistent reporting of either gender or sexual contact.

	FEMALES			MALES		
Birth Cohort	**%**	**s.e.**	**Unwtd. N**	**%**	**s.e.**	**Unwtd. N**
Pre-1920	1.5%	0.63%	548	4.2%	1.16%	303
1920s	1.7%	0.47%	763	3.3%	0.71%	503
1930s	1.9%	0.50%	818	5.1%	0.95%	647
1940s	2.4%	0.44%	1,286	4.4%	0.65%	1,108
1950s	5.3%	0.53%	1,946	5.0%	0.57%	1,551
1960s	5.7%	0.56%	1,898	5.1%	0.57%	1,529
1970s and later	6.5%	0.81%	1,129	4.8%	0.76%	938
	P < 0.0001		*lin, p < 0.0001*	*P = 0.7417*		*lin, p =0.2634*

NOTE. Cases were not included in the database if responses were not available for at least two sexual behavior questions for which there was a logically required consistency of response or if a gender code was not available from the household ennumeration form. See text note 14.