



**Report prepared by**

Kara Joyner

Wendy D. Manning

Krista K.  
Westrick\_Payne

Lee A. Brady

# Contextual Heterosexism Database, Phase 1 User Guide



## Acknowledgments

Add Health is funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

This research was supported by grant R01MD016417 from the National Institute on Minority Health and Health Disparities. Further information may be obtained by contacting Kara Joyner ([kara.joyner@utsa.edu](mailto:kara.joyner@utsa.edu)) or Wendy Manning ([wmanning@bgsu.edu](mailto:wmanning@bgsu.edu)).

Joyner et al. would like to acknowledge the work of this ancillary study's advisory team members Madina Agénor, Christopher Carpenter, Lauren Gaydosh, and Taylor Hargrove as well as project consultants Carolyn T. Halpern and Kathleen Mullan Harris. Their suggestions, guidance, and advice were invaluable.

## Citation

Citations of this Add Health User Guide should use the following format:

*Joyner, K., Manning, W. D., Westrick-Payne, K. K., Brady, L. A. (2023). The National Longitudinal Study of Adolescent to Adult Health (Add Health), Contextual Heterosexism Database - Phase 1 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. <https://doi.org/10.17615/7cyw-wn77>.*

## Introduction

This contextual database, Contextual Heterosexism Database - Phase 1 (CHD1), further expands the collection of contextual data available to users of The National Longitudinal Study of Adolescent to Adult Health (Add Health) through the provision of state, county, and tract level measures from the Decennial Census of Population and Housing, American Community Survey (ACS), the Movement Advancement Project (MAP), Lax and Phillips (2009), Public Religion Research Institute (PRRI), Cooperative Election Study (CES), U.S. Religion Census, and Massachusetts Institute of Technology (MIT) Election Lab. These data include indicators of social policies, social climate, and confounding factors related to the study/measurement of structural heterosexism that correspond to Waves 3, 4, and 5. Some of these indicators are new to the Add Health contextual database and others were previously not available at all three of these waves.

## Subject Index

The Subject Index lists the various topics/scales measured by the variables contained in the Contextual Heterosexism Database - Phase 1 as well as other Add Health Ancillary Studies (dependent upon topic and/or wave) relevant to understanding the effects of structural heterosexism. Subtopics capturing variables that exist in other Add Health Ancillary Studies are indicated in italics. A crosswalk of ancillary studies needed to study heterosexism by subtopic and wave can be found in Table 1 (below). To access codebooks and user's guides associated with each ancillary study outlined in Table 1, please see the following link (<https://addhealth.cpc.unc.edu/documentation/codebooks/>) and click on "Contextual Data Files." Links to each are also provided in each section under [Source Description](#).

1. Policies
  - a. Movement Advancement Project (MAP) Overall Policy Tallies
  - b. Year of Adoption by Policy Type (from MAP)
    - i. Adoption legislation
    - ii. Employment anti-discrimination
    - iii. Hate crime protection
    - iv. Housing anti-discrimination
    - v. Public accommodation anti-discrimination
    - vi. Same-sex union recognition
2. Social Climate
  - a. Proportion approving LGB policies
  - b. Proportion voting for republican presidential candidate
  - c. Religious conservatism
  - d. *Sexual minority concentrations*
3. Income Inequality
  - a. Household income by sex composition of couple
4. Confounding Variables
  - a. Gini index
  - b. *Population size/density/classification*
  - c. *Proportion owner-occupied housing units*
  - d. *Socioeconomic status (SES) scale (5 variables)*
  - e. *RUCA codes*

**TABLE 1: MEASURES OF HETEROSEXISM - ADD HEALTH ANCILLARY STUDY CROSSWALK**

SUBTOPICS	ANCILLARY STUDIES NEEDED		
	W3	W4	W5
<b>Policies</b>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> </ul>
<b>Social climate</b>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> <li>Wave III Contextual Data</li> <li>Wave I, II, III Political Context Data</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> <li>Contextual Wave IV Database,</li> <li>Wave IV Supplemental Tract-Level Contextual Data</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> <li>Contextual Wave V Database</li> </ul>
<b>Income inequality</b>		<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> </ul>
<b>Confounders</b>	<ul style="list-style-type: none"> <li>Wave III Contextual Data</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Wave IV Database,</li> <li>Wave IV Supplemental Tract-Level Contextual Data</li> </ul>	<ul style="list-style-type: none"> <li>Contextual Heterosexism Database - Phase 1</li> <li>Contextual Wave V Database</li> </ul>

## Data Structure and Form

The data file (**w5htxsm1.sas7bdat**) contains a total of 73 variables for Add Health respondents who participated in one or more of Waves III, IV and V. The first variable is the respondent identifier (AID), by which these contextual data can be merged with other Add Health data files. The remaining variables include measures across four subtopics. A list of all variables can be found in the [Data Dictionary](#) section.

The contextual variables in the CHD1 measure characteristics associated with specific locations in the U.S. at varying geographic levels, including state, county, and tract. Add Health participants are linked to these contextual measures based on their residency locations at each conducted interview. To protect participants from identification/disclosure Add Health provides pseudo FIPS codes for respondents' residences at Waves I through V (e.g., Wave V Grouping File, w1\_5grp.xpt). These location identifiers are based on 2010 Census geographic boundaries and are longitudinally consistent across all waves.

## Source Descriptions

### American Community Survey

The American Community Survey is a nationwide survey conducted by the U.S. Census Bureau. The questionnaire is administered every month of every year since 2005 and is an expansion of the Decennial Census asking questions on education, employment, internet access, and transportation (to name a few). The questionnaire is sent to a sample of addresses in the 50 states, District of Columbia, and Puerto Rico at a total of about 3.5 million solicitations a year. The ACS allows for the production of period estimates of the U.S. population and are released as 1-year and 5-year estimates and data products. The 1-year estimates represent 12 months of collected data. Geographic areas with populations of at least 65,000 people are identifiable. ACS data for producing 1-year estimates are available from 2005 and onward – as of this publication the most recent available data represent the year 2022. Beginning in 2016 1-Year Supplemental Estimates have also been published from the ACS. These data also represent 12 months of collected data, are provided as simplified versions of ACS tables, and are published for geographic areas with populations of at least 20,000 people. The 1-year Supplemental Estimates data are available from 2014 and onward – as of this publication the most recent available data represent the year 2022. While the 1-year estimates provide the most current data they have larger margins of error than the 5-year estimates. Beginning in 2010, 5-year estimates, representing 60 months of collected ACS data have been released for all geographic areas regardless of population size. The 5-year estimate data are available from 2005-2009 and onward. Several measures for this database were derived from ACS estimates, including:

#### Measures of Social Climate

1. **Proportion of households headed by different sex unmarried partners among all households** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
2. **Proportion of households headed by same sex unmarried partners among all households** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
3. **Proportion of households headed by male couples of the same sex among all households** (included in this study – Contextual Heterosexism Database - Phase 1)
4. **Proportion of households headed by female couples of the same sex among all households** (included in this study – Contextual Heterosexism Database - Phase 1)

#### Measures of Inequality

5. **Median household income** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
6. **Proportion of households below the poverty level** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
7. **Unemployment rate of males 16 years and over** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
8. **Proportion 25 years or older with a bachelor's degree or more** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))

9. **Proportion of employed 16 years and over in management, professional, and related occupations** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
10. **Gini Coefficient – dispersion of income across the entire income distribution** (included in this study – Contextual Heterosexism Database - Phase 1)
11. **Median household income for male couples / median household income for different sex couples** (included in this study – Contextual Heterosexism Database - Phase 1)
12. **Median household income for female couples / median household income for different sex couples** (included in this study – Contextual Heterosexism Database - Phase 1)
13. **Median household income for different sex couples** (included in this study – Contextual Heterosexism Database - Phase 1)

#### Measure of Residential Instability

14. **Proportion of owner-occupied homes (units)** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))

#### Measures of Population Size and Density

15. **Total population** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))
16. **Density—person per square km** (included in the following Add Health Ancillary Studies: [Contextual Wave IV Database](#), [Contextual Wave V Database](#) and [Wave IV Supplemental Tract-Level Contextual Data](#))

The majority of the 16 measures identified above are available from multiple different Add Health Ancillary Studies. The different studies are noted with their associated documentation files hyperlinked in parentheses. Joyner et al. compiled the remaining six measures for the CHD1. They include the proportion of households headed by male couples of the same sex among all households and the proportion of households headed by female couples of the same sex among all households. These were derived from Census Table B11009: *Coupled Households by Type* from ACS 5-Year Estimates Detailed Tables. The state-level median household income measures were derived in Stata/SE 16.0 by Joyner et al. using 2009 and 2018 ACS 5-year data files from the IPUMS USA database (Ruggles, et al., 2023). Finally, the Gini Index is from Census Table B19083: *Gini Index of Income Inequality* from ACS 5-Year Estimates Detailed Tables. It is important to note that the measurement and reporting of sexual minorities by the U.S. Census Bureau with U.S. Decennial Census and American Community Survey data have evolved over time. As such, concerns regarding measurement reliability and validity over time and across data collections have been raised. See the Data Note in the Appendix for more details.

The following table presents an index of the American Community Survey estimates used in the generation of the contextual variables comprising the Contextual Heterosexism Database - Phase 1 file.

**TABLE 2: INDEX OF 2009 5-YEAR ESTIMATES FROM TABLE B11009: COUPLED HOUSEHOLDS BY TYPE & 2018 5-YEAR ESTIMATES FROM IPUMS-USA**

Source Variable Name	Description
<b>B11009_001E</b>	Total households
<b>B11009_003E</b>	Unmarried-partner households - male householder and male partner
<b>B11009_005E</b>	Unmarried-partner households - female householder and female partner
<b>B19083_001E</b>	Gini index
<b>GQ</b>	Group quarters status
<b>HHINCOME</b>	Total household income
<b>HHWT</b>	Housing weight
<b>RELATED</b>	Relationship to household head
<b>SERIAL</b>	Household serial number
<b>SEX</b>	Sex
<b>STATEFIP</b>	State (FIPS code)
<b>YEAR</b>	Census year

#### Data Citations

Gaydosch, L., Hargrove, T., Dennis, A. C., Frizzelle, B., & Horowitz, J. (2021). The National Longitudinal Study of Adolescent to Adult Health (Add Health), Contextual Wave V Database [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. [https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted\\_use/ContextualWaveVDatabase.zip](https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted_use/ContextualWaveVDatabase.zip)

Hargrove, T., Gaydosch, L. & Dennis, A. C. (2021). The National Longitudinal Study of Adolescent to Adult Health (Add Health), Contextual Wave IV Database [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. [https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted\\_use/ContextualWaveIVDatabase.zip](https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted_use/ContextualWaveIVDatabase.zip)

Morales, L & Monbureau, T. (2013). The National Longitudinal Study of Adolescent to Adult Health (Add Health), Contextual Wave IV Supplemental Tract-Level Contextual Data [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. [https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted\\_use/TRACT4\\_new.zip](https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted_use/TRACT4_new.zip)

Ruggles, S., Flood, S., Sobek, M., Brockman D., Cooper, G., Richards, S., Schouweiler, M. (2023). IPUMS USA: Version 13.0 [dataset]. Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D010.V12.0>

U.S. Census Bureau. (2009). Unmarried-Partner Households by Sex of Partners [data table B11009 for 2009]. *2005-2009 American Community Survey 5-Year Estimates* (all census tracts, counties, and states). Retrieved from <https://data.census.gov/>

U.S. Census Bureau. (2018). Unmarried-Partner Households by Sex of Partners [data table B11009 for 2009]. *2014-2018 American Community Survey 5-Year Estimates* (all census tracts, counties, and states). Retrieved from <https://data.census.gov/>

## Decennial Census

The Constitution of the United States, Article 1, Sections 2 and 9 directs that a census (or enumeration) of the U.S. population be taken. Administered by the US. Census Bureau every 10 years, the data collected are used to apportion the number of seats each state has in the U.S. House of Representatives. The exact questions asked has varied over the years. For the purposes here, Census data were used to provide measures of social climate, including:

1. **Proportion of households headed by different sex unmarried partners among all households**  
(included in Add Health Ancillary Study: [Wave III Contextual Data](#))
2. **Proportion of households headed by same sex unmarried partners among all households**  
(included in Add Health Ancillary Study: [Wave III Contextual Data](#))
3. **Proportion of households headed by male couples of the same sex among all households**  
(included in this study – Contextual Heterosexism Database - Phase 1)
4. **Proportion of households headed by female couples of the same sex among all households**  
(included in this study – Contextual Heterosexism Database - Phase 1)

Proportions 1 and 2 (above) are available in the Add Health Ancillary Study Wave III Contextual Data file (Swisher, 2008). Proportions 4 and 5 (above) were derived from Census Table PCT001: *Unmarried-Partner Households by Sex of Partners* from DEC Summary File 3. It is important to note that the measurement and reporting of sexual minorities by the U.S. Census Bureau with U.S. Decennial Census and American Community Survey data have evolved over time. As such, concerns regarding measurement reliability and validity over time and across data collections have been raised. See the Data Note in the [Appendix](#) for more details.

As indicated above, additional items from the 2000 Decennial Census that may be useful in the study of heterosexism were also archived and merged by Raymond Swisher (Add Health Ancillary Study: [Wave III Contextual Data](#)), including:

1. SES Scale:
  - a. Median household income
  - b. Proportion of households below the poverty level
  - c. Unemployment rate of males 16 years and over
  - d. Proportion 25 years or older with a bachelor’s degree or more
  - e. Proportion of employed 16 years and over in management, professional, and related occupations
2. Residential Instability:
  - a. Proportion of owner-occupied homes (units)
3. Population Size and Density:
  - a. Total population
  - b. Density—person per square km

The following table presents an index of the 2000 Decennial Census estimates used in the generation of the contextual variables comprising the CHD1 file.

**TABLE 3: INDEX OF 2000 DECENNIAL CENSUS ESTIMATES FROM TABLE PCT001: UNMARRIED-PARTNER HOUSEHOLDS BY SEX OF PARTNERS**

Source Variable Name	Description
PCT001001	Total households
PCT001003	Total unmarried-partner households - male householder and male partner
PCT001005	Total unmarried-partner households - female householder and female partner

Data Citations

Swisher, R. R. (2008). The National Longitudinal Study of Adolescent to Adult Health (Add Health), Wave III Contextual Database [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. [https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted\\_use/Wave-III-Contextual-Data.zip](https://addhealth.cpc.unc.edu/wp-content/uploads/docs/restricted_use/Wave-III-Contextual-Data.zip)

U.S. Census Bureau. (2000). Unmarried-Partner Households by Sex of Partners [data table PCT001 for 2000]. *2000 U.S. Decennial Census, Summary File 3* (all census tracts, counties, and states). Retrieved from <https://data.census.gov/>

## Harvard’s Cooperative Election Study (formerly the Cooperative Congressional Election Study) (CCES)

The CCES is a 50,000+ person national stratified sample survey administered by YouGov. In election years the survey consists of two waves—a pre- and post-election wave. In the pre-election wave, respondents answer two-thirds of the questionnaire. This segment of the survey asks about general political attitudes, various demographic factors, assessment of roll call choices, political information, and vote intentions. The pre-election wave is in the field from late September to late October. In the post-election wave, respondents answer the other third of the questionnaire, mostly consisting of items related to the election that just occurred. The post-election wave is administered in November. In non-election years there is a single wave conducted in the fall. For our purposes we relied upon data from election years only due to small sample sizes in non-election years.

The CCES survey question about LGBTQ+ rights included here was in regard to same-sex marriage. The question wording has changed slightly over the years as outlined below.

**TABLE 4: INDEX OF COOPERATIVE ELECTION SURVEY QUESTION WORDING REGARDING VIEWS ON SAME-SEX MARRIAGE LEGISLATION**

Data Collection Year	Question wording
<b>2006</b>	President Bush recently spoke out in favor of a Constitutional Amendment defining marriage as strictly between a man and a woman. Do you support or oppose a Constitutional amendment banning gay marriage?
<b>2007</b>	President Bush recently spoke out in favor of a Constitutional Amendment defining marriage as strictly between a man and a woman. Do you support or oppose a Constitutional amendment banning gay marriage?
<b>2008</b>	Congress considered many important bills over the past two years. For each of the following tell us whether you support or oppose the legislation in principle. Constitutional Amendment banning Gay Marriage.
<b>2014</b>	Do you favor or oppose allowing gays and lesbians to marry legally?
<b>2015</b>	Do you favor or oppose allowing gays and lesbians to marry legally?
<b>2016</b>	Do you favor or oppose allowing gays and lesbians to marry legally?

The following table presents an index of the Cooperative Election Study variables used in the generation of the contextual variables comprising the CHD1 file. Description text aligns with descriptions provided in CCES codebooks for each survey year.

**TABLE 5: INDEX OF COOPERATIVE ELECTION STUDY**

Source Variable Name	Year	Description
<b>V1002</b>	2006	State of residence
<b>V1004</b>	2006	FIPS code
<b>V2103</b>	2006	Gay Marriage Amendment
<b>V206</b>	2008	State
<b>V251</b>	2008	statefips
<b>CC316F</b>	2008	Roll Call Votes - Amendment to Ban Gay Marriage
<b>INPUTSTATE</b>	2014	Pre-Election State Name
<b>CC14 327</b>	2014	Gay Marriage
<b>INPUTSTATE</b>	2016	Pre-Election State Name
<b>CC16 335</b>	2016	Gay Marriage

Data Citations

Ansolabehere, S. (2010). Cooperative Congressional Election Study Common Content, 2006 (Version 4) [dataset]. M.I.T. <https://doi.org/10.7910/DVN/Q8HC9N>

Ansolabehere, S. (2010). Cooperative Congressional Election Study Common Content, 2008 (Version 6) [dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/YUYIVB>

Ansolabehere, S. & Schaffner, B. F. (2017). Cooperative Congressional Election Study Common Content, 2014 (Version 1) [dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/XFXJYY>

Ansolabehere, S. & Schaffner, B. F. (2017). Cooperative Congressional Election Study Common Content, 2016 (Version 4) [dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/GDF6Z0>

## Lax & Phillips

In a 2009 publication Lax and Phillips investigated the effects of policy-specific public opinion on the adoption of state-level policies that are supportive of gay, lesbian and homosexual individuals, the term used in the Lax and Phillips paper is “pro-gay policy support.” We note that these items referenced homosexual as well as both gay and lesbian individuals. For estimating state-level support of “pro-gay” policies, the researchers used data from forty-one national polls compiled in the Roper Center’s iPoll archive. The original polls were random national samples conducted by Gallup, Pew, ABC News, CBS News, AP, Kaiser, and *Newsweek*. The polls yielded approximately 80,000 responses regarding various policies over a ten-year period (1999-2008). The above-mentioned polls asked questions on the following policy types with the general wording of the questions asked as follows:

**TABLE 6: INDEX OF GENERAL SURVEY QUESTION WORDING REGARDING VIEWS ON “PRO-GAY” POLICY ISSUES**

Policy Type	General Question Wording
<b>Adoption</b>	Do you think there should be adoption rights for gay and lesbian couples?
<b>Hate crimes</b>	If a hate crime law were enacted in your state, do you think that homosexuals should covered?
<b>Health</b>	Should there be health insurance and other employee benefits for gay spouses?
<b>Housing</b>	Should there be laws protecting homosexuals from discrimination in housing?
<b>Jobs</b>	Should there be laws to protect gays and lesbians from discrimination in job opportunities?
<b>Marriage</b>	Do you favor allowing gay and lesbian couples to marry legally?
<b>Sodomy</b>	Do you think homosexual relations between consenting adults should be legal?
<b>Unions</b>	Do you favor allowing gay and lesbian couples to form legally recognized civil unions, giving them many of the legal rights of married couples?

The support for each policy was estimated via multilevel regression and poststratification (MRP) to produce one single internally consistent data set. The first stage estimated individual responses to each policy as a function of demographic (sex, race, age, and educational attainment) and geographic (percent evangelical Protestants and Mormons, percent voting Democrat in 2004 presidential election) predictors. State-level predictors for each respondent’s state of residence were used to estimate the state-level effects. Next, estimates for each demographic-geographic respondent type were weighted (post-stratified) by the percentage of each type in the actual state populations. The result was the creation of new estimates of state-level support for the above-mentioned eight policies. We took estimates of explicit pro-gay policy support by state directly from Table 1 *Opinion Estimates and Summary Statistics* (p. 373). For explicit details on the estimation method, please see the Appendix in the Lax and Phillips (2009) publication. We note these attitudes cover a 10-year time span from 1998-2008.

Data Citation

Lax, J. & Phillips, J. (2009). Gay rights in the states: Public opinion and policy responsiveness. *American Political Science Review*, 103(3), 367 - 386. <https://doi.org/10.1017/S0003055409990050>

## Movement Advancement Project (MAP)

Data from MAP were retrieved from [lgbtmap.org](http://lgbtmap.org), backlog data were retrieved via Internet Archive (<https://archive.org/web/>). Two types of variables were derived from these data. The first type is based on years of adoption of specific categories of laws/policies. The year indicates the earliest policy adoption (i.e., if a law was passed in 2005, went into effect in 2006, and was amended in 2007, the year reported is 2006). For Wave III (2011-2002) MAP year variables (SMAP3001-SMAP3006) the value is the year the law/policy was enacted if enacted before 2003, otherwise missing. Regarding Wave IV (2007-2008) MAP year variables (SMAP4001-SMAP4006) the value is the year the law/policy was enacted if enacted before 2009, otherwise missing. Finally, for Wave V (2016-2018) MAP year variables (SMAP5001-SMAP5006) the value is the year the law/policy was enacted if enacted before 2019, otherwise missing. The laws/policies included were three separate measures of non-discrimination including public accommodation anti-discrimination, housing anti-discrimination, and employment anti-discrimination. There was one criminal justice law/policy measure—hate crime protection. Finally, there were two laws/policies related to parental and relationship recognition including adoption legislation and marriage/civil union recognition. MAP year variables are adapted from author-archived backlogs of web-based data visualizations, and thus do not have original variable names associated with them.

1. Nondiscrimination
  - a. Public accommodation anti-discrimination
  - b. Housing anti-discrimination
  - c. Employment anti-discrimination
2. Criminal Justice
  - a. Hate crime protection
3. Parental and relationship recognition
  - a. Adoption legislation
  - b. Marriage/Civil union

The second type of included variables from MAP are tallies based on over 50 different LGBTQ-related laws and policies tracked by MAP. As defined by MAP, “a state’s policy tally scores the laws and policies within each state that shape LGTB people’s lives, experiences, and equality.” The main categories captured include: Relationship & Parental Recognition, Nondiscrimination, Religious Exemptions, LGBTQ Youth, Health Care, Criminal Justice, and Identity Documents. The “overall policy tallies” are distinct from specific sexual orientation or gender identity tallies. Here we provide variables indicating individual state overall policy tallies (SOGI) at Wave V (SMAP5007) and those specific to sexual orientation (SO) at Wave V (SMAP5008). The SO tally includes laws that explicitly mention sexual orientation. For these purposes “sexual orientation” is loosely defined as “a person’s pattern of emotional, romantic, or sexual attraction (or lack thereof) to people.” Lower values on the tallies indicate more or more aggressive anti-LGBTQ+ policies or weaker or fewer pro-LGBTQ+ policies. Higher values indicate fewer or less aggressive anti-LGBTQ+ or stronger/more pro-LGBTQ+ policies.

4. Policy Tallies
  - a. Overall policy tally
  - b. Sexual orientation tally

**Note:** Individual users can define gender identity tallies (GI) by subtracting the SO tally value from the SOGI tally value. The GI tally tracks laws that explicitly address or impact gender identity and/or expression. For these purposes “gender identity” is defined as “a person’s deeply-felt inner sense of being

male, female, or another gender(s)” whereas “gender expression” is in references to “a person’s characteristics and behaviors such as appearance, dress, mannerisms, and speech patterns, all of which can be described as masculine, feminine, or something else.” Gender identity and expression are independent of sexual orientation.

#### Data Citation

Movement Advancement Project. "Equality Maps Snapshot: LGBTQ Equality By State."  
<https://www.mapresearch.org/equality-maps/>. Accessed via Internet Archive. <https://archive.org/web/>.

## MIT Election Lab

The MIT Election Data and Science Lab collects, analyzes, and shares data related to elections and how they are conducted. From their data offerings, we have appended variables derived from their County Presidential Election Returns 2000-2020 data set which contains county-level returns for presidential elections from 2000 to 2020 and their U.S. President 1976-2020 data set which contains constituency (state-level) returns for elections to the U.S. presidency.

The following table presents an index of the County Presidential Election Returns 2000-2020 variables used in the generation of the contextual variables comprising the CHD1 file.

**TABLE 7: INDEX OF COUNTY PRESIDENTIAL ELECTION RETURNS 2000-2020**

Source Variable Name	Description
<b>YEAR</b>	Election year
<b>COUNTY_FIPS</b>	County FIPS code
<b>PARTY</b>	Party of the candidate
<b>CANDIDATEVOTES</b>	Votes received by this candidate for this particular party
<b>TOTALVOTES</b>	Total number of votes cast in this county-year

The following table presents an index of the U.S. President 1976–2020 variables used in the generation of the contextual variables comprising the CHD1 file.

**TABLE 8: INDEX OF U.S. PRESIDENT 1976–2020**

Source Variable Name	Description
<b>YEAR</b>	Election year
<b>STATE_FIPS</b>	State FIPS code
<b>PARTY_SIMPLIFIED</b>	Party of the candidate
<b>CANDIDATEVOTES</b>	Votes received by this candidate for this particular party
<b>TOTALVOTES</b>	Total number of votes cast in this state-year

## Data Citations

MIT Election Data and Science Lab (2018). "County Presidential Election Returns 2000-2020." Harvard Dataverse, V11, UNF:6:HaZ8GWG8D2abLleXN3uEig== [fileUNF]. <https://doi.org/10.7910/DVN/VOQCHO>

MIT Election Data and Science Lab (2017). "U.S. President 1976–2020." Harvard Dataverse, V7, UNF:6:MkQHx147hJcGscG5IqK77g== [fileUNF]. <https://doi.org/10.7910/DVN/42MVDX>

## PRRI American Values Atlas

PRRI is a nonprofit, nonpartisan organization that focuses on the intersection of religion, values, and public life (<https://ava.prri.org/prri>) and publishes The America Values Atlas (AVA) from which we pulled state-level measures of LGB climate (i.e., proportion in favor of LGB policies) in 2015 (see Table 9 for policy type and question wording) from their on-line dashboard. Data collection for the AVA is based on stratified, single-stage, random-digit-dialing (RDD) sample of landline telephone households and randomly generated cell phone numbers. The sample is designed to represent the total U.S. adult population (aged 18 and older) from all 50 states. The landline and cell phone samples are provided by Marketing Systems Group. The AVA conducts at least 50,000 interviews a year over all 52 weeks with about 1,000 interviews collected each week. The resulting sample size is larger than an average public opinion survey. Specific to the data collected in 2015, results for questions on LGBTQ issues are based on a subset of 42,586 telephone interviews conducted between April 29, 2015, and January 7, 2016. A complete roster of state-level sample sizes for 2015 can be found here: <https://ava.prri.org/methodology-2015>. PRRI applies weights calibrated to the American Community Survey estimates for each year via an iterative proportional fitting (IPF) process. In 2015 weights are also trimmed so that they do not exceed 4.0 or fall below 0.25 to prevent single interviews from having too much influence on final results. The following table presents an index of the questions from PRRI American Values Atlas used to create state-level contextual variables for the CHD1 file.

Work and guidance on the PRRI American Values Atlas from Madeline Smith-Johnson was instrumental in the development of these measures.

Zhe, Z, Smith-Johnson, M, & Tumin, D. (2023). Contextual influences on nonresponse to health survey questions about sexual orientation and gender identity. *LGBT Health*.  
<https://doi.org/10.1089/lgbt.2022.0320>

**TABLE 9: INDEX OF QUESTIONS FROM PRRI AMERICAN VALUES ATLAS**

<b>Policy Type</b>	<b>QUESTION HEADER: All in all, do you strongly favor, favor, oppose, or strongly oppose:</b>
<b>Same-sex marriage</b>	Allowing gay and lesbian couples to marry legally?
<b>LGBTQ nondiscrimination laws</b>	Laws that would protect gay, lesbian, bisexual, and transgender people against discrimination in jobs, public accommodations, and housing?
<b>Religiously based refusals to serve gay and lesbian people</b>	Allowing a small business owner in your state to refuse to provide products or services to gay or lesbian people, if doing so violates their religious beliefs?

### Data Citation

PRRI The American Values Atlas (n.d.). LGBTQ State-level Policy Support, 2015. Retrieved from <https://ava.prri.org/#lgbt>

## U.S. Religion Census - Religious Congregations and Membership Study, 2000

Data were obtained from The Association of Religion Data Archives (ARDA) at <https://www.thearda.com/>.

The Religious Congregations and Membership Study, 2000 (RC 2000, for short), was designed and collected by the Association of Statisticians of American Religious Bodies (ASARB). It represents data on 149 religious bodies with the specific number of congregations within each county of the United States. Where available, they also report membership numbers (as defined by the religious body) and total adherents. Participants included 149 Christian denominations, associations, or communions (including Latter-day Saints and Unitarian/Universalist groups); two specially defined groups of independent Christian churches; Jewish and Islamic totals; and counts of temples for six Eastern religions.

While these data contain membership data for many religious groups in the United States, they do not include every group. Users may want to refer to a paper by Roger Finke and Christopher P. Scheitle that explains the "adjusted" adherence rates included in the file.

Finke, R. & Scheitle, C. P. (2005). Accounting for the uncounted: Computing correctives for the 2000 RCMS Data. *Review of Religious Research* 47(1), 5-22.

<https://www.thearda.com/ARDA/archives/Accounting%20for%20the%20Uncounted.pdf>

The following table presents an index of the U.S. Religion Census - Religious Congregations and Membership Study, 2000 variables used in the generation of the contextual variables comprising the CHD1 file.

**TABLE 10: INDEX OF U.S. RELIGION CENSUS - RELIGIOUS CONGREGATIONS AND MEMBERSHIP STUDY, 2000**

Source Variable Name	Description
<b>RELIGIOUS_BODIES</b>	Denomination or group name.
<b>TRADITION</b>	Classification of denominations as Evangelical Protestant, Black Protestant, Mainline Protestant, Orthodox, Catholic, or other.
<b>ADHERENCE_RATE</b>	Members, non-member children, and other regular participants who are not considered members, per 1,000 population.
<b>TOTAL_ADHERENTS</b>	Total number of adherents per geographic unit
<b>STATE</b>	State of congregation
<b>COUNTY</b>	County of congregation

### Data Citation

Association of Statisticians of American Religious Bodies (2002). U.S. Religion Census - Religious Congregations and Membership Study, 2000 (County File) [dataset]. The Association of Religion Data Archives, <https://www.thearda.com/data-archive?fid=RCMSCY>. DOI 10.17605/OSF.IO/AV9KG

## U.S. Religion Census - Religious Congregations and Membership Study, 2010

Data were obtained from The Association of Religion Data Archives (ARDA) at <https://www.thearda.com/>.

The Religious Congregations and Membership Study, 2010 (RC 2010, for short), was designed and collected by the Association of Statisticians of American Religious Bodies (ASARB). It represents data on the number of congregations and adherents for 236 religious groups in each county of the United States. Participants included 217 Christian denominations, associations, or communions (including Latter-day Saints, Messianic Jews, and Unitarian/Universalist groups); counts of Jain, Shinto, Sikh, Tao, and National Spiritualist Association congregations, and counts of congregations and adherents from Baha'is, three Buddhist groupings, four Hindu groupings, four Jewish groupings, Muslims, and Zoroastrians. The 236 groups reported a total of 344,894 congregations with 150,686,156 adherents, comprising 48.8 percent of the total U.S. population of 308,745,538 in 2010.

The following table presents an index of the U.S. Religion Census - Religious Congregations and Membership Study, 2010 variables used in the generation of the contextual variables comprising the CHD1 file.

**TABLE 11: INDEX OF U.S. RELIGION CENSUS - RELIGIOUS CONGREGATIONS AND MEMBERSHIP STUDY, 2010**

Source Variable Name	Description
<b>RELIGIOUS_BODIES</b>	Denomination or group name.
<b>TRADITION</b>	Classification of denominations as Evangelical Protestant, Black Protestant, Mainline Protestant, Orthodox, Catholic, or other.
<b>ADHERENCE_RATE</b>	Members, non-member children, and other regular participants who are not considered members, per 1,000 population.
<b>TOTAL_ADHERENTS</b>	Total number of adherents per geographic unit
<b>STATE</b>	State of congregation
<b>COUNTY</b>	County of congregation

### Data Citation

Grammich, C., Hadaway, K., Houseal, R., Jones, D. E., Krindatch, A., Stanley, R., & Taylor, R. H. (2018, December 11). U.S. Religion Census Religious Congregations and Membership Study, 2010 (County File) [dataset]. The Association of Religion Data Archives, <https://www.thearda.com/data-archive?fid=RCMSCY10> 10.17605/OSF.IO/QUN29

## Variable Naming Conventions

### First character—Geographic level of the variable

Refers to the geographic area to which the variable corresponds. Geographic levels include:

T	Tract
C	County
S	State

Please note, administrative boundaries may have changed between 1990 and 2010. When doing longitudinal analyses, observed change may be, in part, artifacts of changes in boundary delineation. Contextual variables associated with Waves I and II are based on 1990 boundaries, Waves III and IV are based on 2000 boundaries, and Wave V are based on 2010 boundaries.

### Second and third characters – Add Health wave and topic

The second character indicates the Add Health wave. The third character, H, indicates the Heterosexism topic.

### Next 3-6 Characters – Subject

The next set of characters of variable length refer to the subject as defined in the [Data Dictionary](#):

GINI	Gini Coefficient
HHINC	Median Household Income
MAP	Movement Advancement Project (MAP)
PROLGB	Population Approving LGB Policies
RELIG	Religious Congregations and Membership
SEXMIN	Sexual Minority Concentrations
VOTING	Elections and Voting

### Final 3 Characters – Sequential numbering

Each unique variable (e.g., Median Household Income for Male Couples) is enumerated from 1 to N (left padded with zeroes to 3 characters) where N is the number of geography/wave combinations for that particular measure.

**TABLE 12: VARIABLE NAMING STRUCTURE**

Geographic level 1 <sup>st</sup> character	Add Health Wave 2 <sup>nd</sup> character	Topic 3 <sup>rd</sup> character	Data source 3-6 characters	Variable number Final 3 characters
<b>T = Tract</b> <b>C = County</b> <b>S = State</b>	3 = Wave III 4 = Wave IV 5 = Wave V	H	GINI = Gini Coefficient HHINC = Median Household Income MAP = Movement Advancement Project PROLGB = Population Approving LGB Policies RELIG = Religious Congregations & Membership SEXMIN = Sexual Minority Concentrations VOTING = Elections and Voting	001 – 013

## Data Dictionary

The tables below list all the variables comprising this Add Health Contextual Heterosexism Database - Phase 1 (CHD1) organized by data source. The table of [Variables Measuring Proportion of Population of Population Approving of LGB Policies](#) is slightly different. Because the variables measuring the proportion of the population approving LGB policies stem from multiple data sources, we present each measure by the respective data source corresponding to each wave of Add Health for which a measure is available.

Refer to the [Variable Naming Conventions](#) section for more information on what the different components of the variable names in the tables below indicate.

### Variables Based on Census and ACS Five-Year Estimates

#### Sexual Minority Concentrations

Name	Description	Formula
<b>S/C/T 3HSEXMIN001</b>	Proportion of households headed by male couples of the same sex among all households	pct001003/pct001001
<b>S/C/T 4HSEXMIN001</b>	Proportion of households headed by male couples of the same sex among all households	b11009_003e/b11009_001e
<b>S/C/T 5HSEXMIN001</b>	Proportion of households headed by male couples of the same sex among all households	b11009_003e/b11009_001e
<b>S/C/T 3HSEXMIN002</b>	Proportion of households headed by female couples of the same sex among all households	pct001005/pct001001
<b>S/C/T 4HSEXMIN002</b>	Proportion of households headed by female couples of the same sex among all households	b11009_005e/b11009_001e
<b>S/C/T 5HSEXMIN002</b>	Proportion of households headed by female couples of the same sex among all households	b11009_005e/b11009_001e

#### Gini Coefficient

Name	Description	Formula
<b>S/C/T 5HGINI001</b>	Dispersion of income across the entire income distribution	b19083_001e

## Median Household Income

Name	Description	Formula
<b>S4HHHINC001</b>	Median household income for female couples	median of HHINCOME *
<b>S5HHHINC001</b>	Median household income for female couples	median of HHINCOME *
<b>S4HHHINC002</b>	Median household income for male couples	median of HHINCOME *
<b>S5HHHINC002</b>	Median household income for male couples	median of HHINCOME *
<b>S4HHHINC003</b>	Median household income different-sex couples	median of HHINCOME *
<b>S5HHHINC003</b>	Median household income different-sex couples	median of HHINCOME *

\* The variable “Householder couple type” was not available in IPUMS-USA until 2019 and the variable to identify “Same-sex married couple[s]” was not available until 2013. As such, we generated our own variable to identify householder couple type using the variables RELATED and SEX prior to calculating the median household incomes by state.

## Variables Measuring Policies Based on MAP (Policy Years)

Name	Description
<b>S3HMAP001</b>	Year state adopted employment anti-discrimination law(s)/statute(s)
<b>S4HMAP001</b>	
<b>S5HMAP001</b>	
<b>S3HMAP002</b>	Year state adopted public accommodation anti-discrimination law(s)/statute(s)
<b>S4HMAP002</b>	
<b>S5HMAP002</b>	
<b>S3HMAP003</b>	Year state adopted housing anti-discrimination law(s)/statute(s)
<b>S4HMAP003</b>	
<b>S5HMAP003</b>	
<b>S3HMAP004</b>	Year state adopted hate crime protection law(s)/statute(s)
<b>S4HMAP004</b>	
<b>S5HMAP004</b>	
<b>S4HMAP005</b>	Year state adopted marriage/civil union protection law(s)/statute(s)
<b>S5HMAP005</b>	
<b>S4HMAP006</b>	Year state adopted adoption protection law(s)/statute(s)
<b>S5HMAP006</b>	
<b>S5HMAP007</b>	MAP SOGI Tallies
<b>S5HMAP008</b>	MAP SO Tallies

## Variables Based on Data from the MIT Elections and Voting Lab

Name	Description	Formula
<b>S/C 4HVOTING001</b>	Proportion of votes cast for the Republican presidential candidate, 2004	$(\text{candidatevotes}/\text{totalvotes})$ if $\text{year} == 2004$ & $\text{partyaff}^* == 5$
<b>S/C 4HVOTING002</b>	Proportion of votes cast for the Republican presidential candidate, 2008	$(\text{candidatevotes}/\text{totalvotes})$ if $\text{year} == 2008$ & $\text{partyaff}^* == 5$
<b>S/C 5HVOTING003</b>	Proportion of votes cast for the Republican presidential candidate, 2012	$(\text{candidatevotes}/\text{totalvotes})$ if $\text{year} == 2012$ & $\text{partyaff}^* == 5$
<b>S/C 5HVOTING004</b>	Proportion of votes cast for the Republican presidential candidate, 2016	$(\text{candidatevotes}/\text{totalvotes})$ if $\text{year} == 2016$ & $\text{partyaff}^* == 5$

\* The original data on party affiliation from MIT Elections and Voting Lab was in string format. We created a new categorical variable “partyaff” with the following values (1) Democrat, (2) Green, (3) Libertarian, (4) Other, and (5) Republican. Further, the original data were in long format with each party candidate per county and year on its own line. Following recoding it was reshaped into wide format with each county on its own line.

## Variables Based on Data from the U.S. Religion Census - Religious Congregations and Membership Study

Name	Description
<b>S/C 3HRELIG001</b>	Rate (per 1,000) of traditional (evangelical Protestant and Mormon) religious adherence
<b>S/C 4HRELIG001</b>	
<b>S/C 5HRELIG001</b>	

## Variables Measuring Proportion of Population Approving of LGB Policies

Name	Policy Type	Description	Source
<b>S4HPROLGB001</b>	Second-parent adoption	Opinion estimates via multilevel regression and poststratification of data from 1999-2008	Lax & Phillips, 2009 *
<b>S4HPROLGB002</b>	Hate crimes		
<b>S4HPROLGB003</b>	Health benefits		
<b>S4HPROLGB004</b>	Housing		
<b>S4HPROLGB005</b>	Jobs		
<b>S4HPROLGB006</b>	Sodomy		
<b>S4HPROLGB007</b>	Civil Unions		
<b>S4HPROLGB008</b>	Marriage <sup>§</sup>		
<b>S4HPROLGB009</b>	Marriage <sup>§</sup>	Two-year mean share of population approving, 2006 & 2008	CCES
<b>S5HPROLGB010</b>	Marriage <sup>§</sup>	Proportion approving, 2015	PRRI
<b>S5HPROLGB011</b>	Marriage <sup>§</sup>	Two-year mean share of population approving, 2014 & 2016	CCES
<b>S5HPROLGB012</b>	LGBTQ Nondiscrimination laws	Proportion approving, 2015	PRRI
<b>S5HPROLGB013</b>	Religious-based refusals to serve	Proportion approving, 2015	PRRI

\* Opinion estimates were taken directly from Table 1 (Lax & Phillips, 2009). See the [Data Source description](#) for more details.

<sup>§</sup> Regarding marriage, there are multiple sources for measuring the proportion of a population approving of policies/laws supporting marriage among couples of the same-sex.

## Missing codes

This dataset has three different codes that indicate the source of missing data.

- -9992 Missing in the source data
- -9991 Respondent was not interviewed in that wave
- -9990 Respondent lacks the geocode necessary for merging the source data

## Appendix – Source Data Notes

### Measurement of Sexual Minorities in U.S. Census Bureau Data

Starting in 2000 for the Decennial Census (DC) and 2008 for the American Community Survey (ACS), sex was assumed to be correctly identified (prior to this, it was assumed “same sex households” had respondents who incorrectly identified household members sex when same-sex married or cohabiting couples were identified in the data), and same-sex married couples were reclassified as same-sex cohabiting couples (Cohn 2011; Gates 2010; Lofquist and Ellis 2011). Numerous researchers identified significant measurement error based on this strategy (Black et al. 2007; Gates 2015; Gates and Steinberger 2009; Kreider and Lofquist 2015; O’Connell and Feliz 2011; O’Connell and Gooding 2006) highlighting that relatively few errors in a large population of different-sex married couples had a substantial impact on the estimates of the relatively small population of same-sex married couples. Recognition of these errors lead to the release of preferred DC estimates, and Gates (2015) modified ACS counts based on allocation flags for sex, marital status, and marriage year.

### References

- Black, D., Gates, G. J., Sanders, S., & Taylor, L. (2007). Demographics of the gay and lesbian population in the United States: Evidence from available systematic data sources. *Demography*, 37, 139–154.
- Cohn, D. V. (2011). *How accurate are counts of same-sex couples?* (Social & Demographic Trends report). Washington, DC: Pew Research Center. Retrieved from <https://www.pewsocialtrends.org/2011/08/25/how-accurate-are-counts-of-same-sex-couples/>
- Gates, G. J. (2010). Same-sex couples in U.S. Census Bureau data: Who gets counted and why (Report). Los Angeles: The Williams Institute, UCLA School of Law. Retrieved from <https://williamsinstitute.law.ucla.edu/wp-content/uploads/SS-Couples-US-Census-Bureau-Data-Aug-2010.pdf>
- Gates, G. J. (2015). Demographics of married and unmarried same-sex couples: Analyses of the 2013 American Community Survey (Report). Los Angeles: The Williams Institute, UCLA School of Law. Retrieved from <http://williamsinstitute.law.ucla.edu/wp-content/uploads/Demographics-Same-Sex-Couples-ACS2013-March-2015.pdf>
- Gates, G. J., & Steinberger, M. (2009, April). *Same-sex unmarried partner couples in the American Community Survey: The role of misreporting, miscoding, and misallocation*. Paper presented at the annual meeting of the Population Association of America, Detroit, MI.
- Kreider, R. M., & Lofquist, D. A. (2015). *Matching survey data with administrative records to evaluate reports of same-sex married couple households* (SEHSD Working Paper No. 2014-36). Washington, DC: U.S. Census Bureau.
- Lofquist, D. A., & Ellis, R. (2011, March). *Comparisons of estimates of same-sex couple households from the ACS and CPS*. Poster presented at the annual meeting of the Population Association of America, Washington, DC.
- O’Connell, M., & Feliz, S. (2011). *Same-sex couple household statistics from the 2010 census* (SEHSD Working Paper No. 2011-26). Washington, DC: U.S. Census Bureau.

O'Connell, M., & Gooding, G. E. (2006, March). *The use of first names to evaluate reports of gender and its effect on the distribution of married and unmarried couple households*. Poster presented at the annual meeting of the Population Association of America, Los Angeles, CA.