Methodology of Estimating Sexual and Gender Identity Population Quantities for Calibration of Nonprobability Survey Data: Using the U.S. Census Bureau's Household Pulse Survey (2021-2024)

Last updated: August 11, 2026

1. Overview and Background

The Human Rights Campaign (HRC) Foundation's LGBTQ+ research panel, Community Insights (CI), is the longest-running and largest opt-in LGBTQ+ panel in the United States. The panel is nonprobability-based and used to conduct timely, high-quality survey research on the issues shaping LGBTQ+ adults lives today across a range of topics such as health equity and workplace inclusion to civic engagement and how institutions and brands can better serve and include the community. Reliable LGBTQ+ data helps organizations, advocates, and policymakers better understand community needs, inform inclusive decision-making, and advance LGBTQ+ equality. One of the critical elements of CI methodology is having benchmarks of the LGBTQ+ adult population for calibration. Calibration using post-stratification or raking is and has been a common practice for nonprobability survey estimation (Baker et al., 2013). The methodology of calibration developed out of probability theory and sampling (Deville & Sarndal, 1992). While able to correct for some coverage error associated with nonprobability sampling, calibration is known to be unable to fully eliminate such bias (Dever, Rafferty, & Valliant, 2008; Tourangeau, Conrad, & Couper, 2013).

In 2020, during the COVID-19 pandemic and prior to acquisition of CI, HRC Foundation partnered with PSB Insights to conduct a series of polls to assess the various ways the pandemic was impacting the LGBTQ+ community. At the time, population benchmarks for ensuring subsamples and oversamples of LGBTQ+ adults were unavailable using traditional sources like the Decennial Census. HRC Foundation utilized available analysis from the William's Institute's LGBT Data and Demographics resource, which was based on 2016 data from the Gallup Daily Tracker. This specific resource is no longer available online, and some elements were later updated with future releases using 2020-2021 Behavior Risk Factor Surveillance System data (see Flores & Conron, 2023). Administered by the Census Bureau, the Census provides an accurate count of how many people live in the United States, as well as demographic data about the United States population. Unfortunately, the Census and many other sources of demographic data do not currently include accurate counts of LGBTQ+ people. However, in July of 2021, the Household Pulse Survey (HPS), a web-based, national household probability survey administered by the Census Bureau became the first in its history to include sexual and gender identity (SOGI) measures in a questionnaire.

Though HPS has its limitations and was labeled experimental, the HPS represents one of the largest federal data collection efforts to have routinely measure SOGI among U.S. adults. The HPS SOGI measures were developed through an interagency process informed by emerging best practices for federal data collection and have subsequently been examined in peer-reviewed

research. Research suggests that responses to these measures exhibit meaningful demographic patterning rather than random measurement error, supporting their potential utility for population estimation and surveillance purposes (Julian et al., 2024; Pao et al., 2025). HRC Foundation pooled four years (2021-2024) of HPS data for a sample of more than 2.2 million U.S. adults. The combined dataset spans weeks 34 to 63 and cycles 1-7 in phase 4.0 and 4.1, providing precision and stability of benchmark estimates for CI's calibration needs. In this technical report,

This technical report proceeds as follows. First, it describes the construction of the pooled HPS analytical file, treatment of item nonresponse and imputed values, and operationalization of SOGI measures. Then it presents the estimation strategy, including construction of pooled person and replicate weights and the procedures used to estimate population proportions, totals, and variances. Next, the report provides the resulting national SOGI benchmark estimates and demographic distributions in tables. It also has a discussion portion regarding the rationale for selecting the HPS, limitations of the data and analysis, and alternative data sources that may be considered for future benchmark development. The final section summarizes the implications of the estimates for calibration of the Community Insights panel. An appendix of distributions of demographic groups by SOGI is also provided at the end of the methodological report.

2. Data

2.1. Analytical File Construction

To construct a pooled benchmark dataset for post-stratification and weighting analyses, public-use microdata from the United States Census Bureau Household Pulse Survey (HPS) were assembled across two harmonized components. First, weekly HPS public-use files from Weeks 34–63 (2021–2023) were imported and appended in Stata SE Version 17, with each weekly respondent-level file merged one-to-one to its corresponding replicate weight file using respondent identifiers `scram` and `week`. Second, 2024 cross-sectional HPS files from Phase 4.0 (Cycles 1–3) and 4.1 (Cycles 4–7) were similarly imported, merged with their corresponding replicate weight files using `scram` and `cycle` and appended into a pooled cross-sectional file. The 4.2 phase was intentionally excluded to preserve questionnaire comparability for SOGI harmonization. Metadata variables identifying source year, week, phase, folder, and file origin were generated to preserve provenance across pooled components. In 2021–2023, the HPS consisted of repeated weekly cross sections. Beginning in Phase 4.0 (2024), the HPS transitioned to approximately monthly cross-sectional collection cycles. The weekly and cross-sectional pooled files were then appended into a unified respondent-level dataset containing 2,452,769 observations (**Table 1**), however, this is not the final sample size used in later analyses as imputed and missing data will be handled and maximum sample size is used for each quantity estimate. Replicate weights (80 person-level and household-level replicate weights), along with main survey weights, were retained to support design-consistent variance estimation. A structured diagnostic protocol was conducted to verify data integrity, including confirmation of expected week and cycle inclusion, checks for excluded phases, duplicate detection on respondent identifiers within source files and waves, verification of successful retention and nonmissingness of all replicate weights, agreement checks between

source-week and survey week indicators, and summary checks of weight distributions across pooled components. Diagnostics indicated no duplicated records, no failed replicate-weight merges, no phase contamination, and complete preservation of design variables, yielding a quality-controlled pooled HPS benchmark file for subsequent harmonization and weighting analyses.

Table 1: Year, Cross-sections, and Sample Sizes within Household Pulse Survey Pooled File		
Year	**Cross-section (week/cycle) within-year**	**Individuals within cross-section, within-year**
2021	Week 34	64,562
	Week 35	68,799
	Week 36	69,114
	Week 37	63,536
	Week 38	59,833
	Week 39	57,064
Total	**6**	**382,908**
2022	Week 40	60,826
	Week 41	74,995
	Week 42	75,482
	Week 43	84,158
	Week 44	63,769
	Week 45	61,767
	Week 46	62,826
	Week 47	58,304
	Week 48	46,801
	Week 49	50,937
	Week 50	42,040
	Week 51	61,453
	Week 52	70,685
Total	**13**	**814,043**
2023	Week 53	68,504
	Week 54	75,709
	Week 55	72,738
	Week 56	61,927
	Week 57	59,290
	Week 58	64,792
	Week 59	63,802
	Week 60	68,830
	Week 61	68,454
	Week 62	72,839
	Week 63	79,371
Total	**11**	**756,256**
2024	Ph4.0 CY1	68,544
	Ph4.0 CY2	71,152
	Ph4.0 CY3	71,798
	Ph4.1 CY4	73,352
	Ph4.1 CY5	72,162

Total	Ph4.1 CY6 Ph4.1 CY7 **7**	72,125 70,429 **499,562**
Overall	**37**	**2,452,769**

2.2. Item Nonresponse and Handling Imputed Cases

Item-level missing data coded as "question seen but category not selected" (-99) and "missing/did not report" (-88) were treated as item nonresponse rather than substantive response categories. These observations were excluded from the construction of benchmark control distributions. Survey-weighted estimates were therefore calculated among respondents providing valid substantive responses to the relevant variables. Consistent with established survey methodological practice, item nonresponse categories were not incorporated into calibration targets. Instead, their prevalence and correlates were evaluated descriptively to assess the potential for nonresponse bias. Responses of "I don't know" to sexual identity questions were treated separately from item nonresponse. Because these responses represent a substantive expression of uncertainty or questioning identity, they were retained as a distinct category in sensitivity analyses examining their influence on benchmark estimates. Records with missing responses on sex at birth were not supplemented using imputed values for target construction, avoiding incorporation of model-based classifications into the benchmark margins themselves, per recommended by the Census Bureau (United States Census Bureau, 2022). As a result, analyses are limited to respondents with non-imputed sex at birth. The methodology also limited based on other variables (birth year, race/ethnicity, and education) to non-imputed data across all cases. A total of 99,245 cases with imputed data are excluded from estimation.

2.3. SOGI Variable Operationalization

2.3.1. Sexual Identity

Sexual identity was derived from the HPS item asking respondents which category best represented how they think of themselves. Response options included *gay or lesbian*, *straight*, *bisexual*, *something else*, and *I don't know*. Item nonresponse, including respondents who did not answer or skipped the question, as well as respondents selecting *I don't know*, were treated as missing and excluded from benchmark estimation. Two operationalizations of sexual identity were constructed. First, a three-category measure consisting of: (1) gay or lesbian, (2) straight, and (3) bisexual-plus (Bi+) was developed. Respondents selecting *bisexual* or *something else* were combined into a broader Bi+ category. This approach preserved stable cell sizes while recognizing that respondents selecting *something else* may represent sexual identities not fully captured by nonmonosexual sexual orientation categories.

2.3.2. Gender Identity

Gender identity was operationalized using the two-step approach available within the Household Pulse Survey (HPS), which combines information on sex assigned at birth and current gender identity. Respondents first reported the sex assigned to them at birth (male or female) and subsequently reported whether they currently describe themselves as male, female, transgender, or "none of these." Responses indicating item nonresponse were treated as missing and excluded from benchmark estimation. These measures were cross-classified to construct a six-category gender identity variable consisting of: (1) assigned male at birth (AMAB) with current gender male; (2) assigned female at birth (AFAB) with current gender female; (3) AMAB with current gender female; (4) AFAB with current gender male; (5) AMAB with current gender as gender-expansive and (6) AFAB with current gender as gender-expansive. For development of post-stratification benchmarks, these categories were subsequently collapsed into a binary gender identity indicator. Respondents whose reported sex assigned at birth aligned with their current gender identity (AMAB male and AFAB female) were classified as cisgender, whereas respondents whose reported sex assigned at birth differed from their current gender identity, as well as those selecting "none of these" on the current gender identity item, were classified within a broader transgender/gender-expansive category. Although the HPS employs a two-step framework consistent with contemporary recommendations for gender identity measurement, the available response categories do not support reliable differentiation of transgender men, transgender women, and other gender-expansive populations at a level sufficient for benchmark construction. Accordingly, a broader gender minority classification was adopted to preserve conceptual validity while ensuring stable population estimates suitable for calibration and weighting purposes.

2.3.3. Combined SOGI Indicator

A combined indicator was constructed by integrating the sexual identity and gender identity measures described above. Respondents were classified as "SGMs" (sexual and gender minorities) if they were identified as either a sexual minority (gay/lesbian or bisexual+) or gender minority (transgender, gender-expansive). Respondents classified as both heterosexual and cisgender were categorized as "non-SGMs." This combined indicator served as the principal benchmark variable for estimating overall population quantities. Respondents are set to missing if they have sexual and gender identity missing or unknown in both cases. Furthermore, straight respondents with missing or unknown gender identity are set to missing and excluding from analyses. Respondents with known sexual identities are included in composite estimates even if gender identity is missing since inclusion in the SOGI indicator is dependent on being either sexual or gender minority.

2.3.4. Residual SOGI Categories

Residual response categories were retained in the construction of the SOGI measures to preserve representation of respondents whose identities may not be adequately captured by conventional response options. Respondents selecting “something else” for sexual identity were classified within a broader bisexual-plus (Bi+) category, while respondents selecting “none of these” for current gender were classified within the broader transgender/gender-expansive category. These responses were treated as substantive identity categories rather than as item nonresponse. Because inclusion of residual categories could theoretically increase the estimated size of the SGM population, their prevalence was examined descriptively. In the non-imputed analytic sample, 5,400 respondents selected both “something else” for sexual identity and “none of these” for current gender. These respondents comprised 0.257% of the unweighted sample. Thus, respondents identified through both residual SOGI categories constitute a small share of the population and do not account for a substantial portion of the overall SGM benchmark estimate. Residual SOGI categories were retained because, when accompanied by an affirmed sexual or gender identity, sufficient information exists to classify respondents within the broader SGM population. Respondents selecting only residual SOGI categories for both sexual orientation ("something else") and gender identity ("none of these") represented 5,400 respondents (0.257% of the unweighted analytic sample). Consequently, any potential misclassification arising from interpretation of residual categories is unlikely to materially affect the overall SGM benchmark estimate. Instead, inclusion of residual categories would be more likely to influence positively the distribution of respondents within SGM subgroups, particularly the bisexual-plus (Bi+) and transgender/gender-expansive benchmark categories, rather than the aggregate SGM population estimate.

2.3.5. Sample Descriptive Statistics of SOGI Variables

Table 2 presents sample sizes used for point estimation and unweighted descriptive statistics of SOGI variables. The estimates are based on data limited to non-imputed cases.
A total of 2,124,493 observations are used for producing estimates with the composite SOGI indicator. For sexual identity, 2,301,916 observations are used. And for gender identity, 2,148,860 observations are used.

Table 2: Summary Descriptive Statistics of Analytical File by SOGI (Unweighted, not for reporting)

	N	**%**
Non-SGM	1,901,559	89.51%
SGM	222,934	10.49%
Total	**2,124,493**	**100%**

Straight	2,092,575	90.91%
Sexual Minority	209,341	9.09%
Gay or lesbian	*78,867*	*3.43%*
Bisexual+	*130,474*	*5.67%*
Total	**2,301,916**	**100%**
Cisgender	2,115,222	98.43%
Cisgender Men	*896,066*	*41.70%*
Cisgender Women	*1,219,156*	*56.74%*
Gender Minority	33,638	1.57%
Total	**2,148,860**	**100%**
Note: Raw and unweighted estimates. Do not use for reporting and benchmarking.		

3. Estimation Strategy

3.1. Pooled Weight Construction

To construct weights for the pooled analysis, the HPS final person weight and each of the 80 replicate weights were divided by 37, corresponding to the total number of survey weeks and cycles included in the pooled dataset. The resulting pooled person weight was used for point estimation, while the 80 correspondingly rescaled replicate weights were used for BRR variance estimation. Let w_{ic} represent the final pooled full-sample weight for respondent i within cross-section c and $w_{ic}^{(r)}$ represent the final pooled replicate weight for $r = 1,\ 2,\ \ldots,\ 80$ replications. The final weights were constructed as

$$w_{ic} = \frac{w_{ic}^{(0)}}{C} \text{ and } w_{ic}^{(r)} = \frac{w_{ic}^{(r),(0)}}{C}$$

where $w_{ic}^{(0)}$ is the original weight prior to rescaling due to pooling of cross-sections and $C = 37$ total cross-sections. Dividing the number of total cross-sections is appropriate as each HPS file is a separately weighted cross-sectional estimate of the population, regardless of whether its collection window was about two weeks or one month (weeks vs. cycles). For those interested in estimates by year, which are not provided in this methodological report, weights should be rescaled such that each set of weights are divided by the number of cross-section within a given year. Let w_{ict} represent the final pooled full-sample weight for a respondent within a cross-section in year t such that $t = \{2021,\ 2022,\ 2023,\ 2024\}$ then

$$w_{ict} = \frac{w_{ict}^{(0)}}{C_t} \text{ and } w_{ict}^{(r)} = \frac{w_{ict}^{(r),(0)}}{C_t}$$

providing within-year person weights. Conceptually, the overall pooled person weights can be represented by extending the within-year weighting framework across all years, where the total number of pooled cross-sections is simply the sum of the within-year cross-sections, such that

$$w_{ic} = \frac{w_{ict}^{(0)}}{\sum_{t=2021}^{2024} C_t} = \frac{w_{ic}^{(0)}}{C} \text{ and } w_{ic}^{(r)} = \frac{w_{ict}^{(r),(0)}}{\sum_{t=2021}^{2024} C_t} = \frac{w_{ic}^{(r),(0)}}{C}$$

The weighting diagnostics for the pooled full-sample weights and weights among the SGM subsample are summarized in **Table 3**. Compared with the full analytic sample, SGM respondents generally received modestly larger calibration weights, as reflected in higher mean, median, and upper-tail percentiles. However, the most extreme weights observed in the pooled dataset were not concentrated among SGM respondents, suggesting that the largest calibration adjustments were not unique to the SGM analytic subgroup.

Table 3: Weighting Diagnostics Among Full Sample and SGM Subsample (U.S. Household Pulse Survey Pooled Data, 2021-2024)

	Full Sample	SGM Subsample
Mean	100.3	114.7
Median	50.5	52.8
90th percentile	219.7	257.2
95th percentile	341.0	408.6
99th percentile	808.2	986.5
99.9th percentile	2,091.8	2,493.2
Minimum	.6523122	.6948509
Maximum	11,547.3	10,004.4
Note: Weights are unstandardized (nonnormalized).		

To evaluate the variability introduced by the pooled analytic weights, the unequal-weighting design effect was calculated (Kish, 1965). Following Kish, the original weights were first normalized by their mean, and the design effect was estimated as the mean squared relative weight:

$$Deff_w = 1 + CV(w)^2$$

$$= \frac{n\sum_{i=1}^{N} w_i^2}{(\sum_{i=1}^{n} w_i)^2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{w_i}{\bar{w}}\right)^2$$

where the pooled HPS analytic weights exhibited an unequal-weighting design effect of approximately 4.2, indicating that unequal weighting reduced the effective sample size relative to an equally weighted simple random sample by a moderate but noticeable effect size worth documenting. Ultimately, the penalties of unequal weighting are effectively addressed by sheer sample volume and design-consistent replication method of variance estimation described in a later section. Note, Kish's unequal-weighting design effect measures the component of variance inflation attributable solely to unequal weighting under an assumed simple random sample. It is presented

as a diagnostic of weight dispersion only. Variance estimation for all analyses relies on Fay's Balanced Repeated Replication (BRR) using the Census Bureau's replicate weights, which accounts for the full complex sample design, including the effects of weighting and the sample design reflected in the replicate weights.

3.2. Proportion and Variance Estimation

Let $\hat{p}$ represent the estimated population proportion represented by the sample proportion and y_{ic} as an indicator or categorical variable for N respondents. Proportions are estimated by

$$\hat{p} = \frac{\sum_{c=1}^{C}\sum_{i=1}^{N} w_{ic}y_{ic}}{\sum_{c=1}^{C}\sum_{i=1}^{N} w_{ic}}$$

with estimated population total

$$\widehat{t_y} = \sum_{c=1}^{C}\sum_{i=1}^{N_c} w_{ic}y_{ic}$$

such that

$$\hat{p} = \frac{\widehat{t_y}}{\sum_{c=1}^{C}\sum_{i=1}^{N} w_{ic}}$$

The proportion estimator is the average population proportion of U.S. adults possessing each sexual identity, gender identity, or combined SOGI characteristic across the pooled HPS cross-sections included in the analysis.

Variance estimation was performed using balanced repeated replication (BRR) with Fay's adjustment of $k = 0.5$, per Census Bureau guidance. The variance of the proportion $\widehat{Var}(\hat{p})$ is represented by

$$\widehat{Var}(\hat{p}) = \frac{1}{R(1-k)^2}\sum_{r=1}^{R}\left(\hat{p}^{(r)} - \hat{p}\right)^2$$

The data were declared as survey data in Stata SE Version 17 by

```
svyset [pw=pweight_pool], brrweight(pweight1_pool-pweight80_pool) ///
                                vce(brr) fay(.5)
```

4. Tables of Estimates

Table 4: Sexual Identity Estimates of U.S Adults (U.S. Household Pulse Survey Pooled Estimates, 2021-2024)		
	% Population Total	[95% CI]
Straight	**89.66%** 205,495,411	[89.58%-89.73%] [205,258,669, 205,732,153]
Sexual Minority	**10.34%** 23,707,378	[10.27%, 10.42%] [23,535,910, 23,878,847]
Lesbian or gay	**3.27%** 7,489,331	[3.22%, 3.32%] [7,370,554, 7,608,107]
Bisexual+	**7.08%** 16,218,048	[7.01%, 7.14%] [16,068,396, 16,367,699]
Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Population totals exclude respondents with "don't know" selected as their sexual identity. Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files.		

Table 5: Gender Identity Estimates of U.S Adults (U.S. Household Pulse Survey Pooled Estimates, 2021-2024)		
	% Population Total	[95% CI]
Cisgender	**97.66%** 211,022,246	[97.62%, 97.70%] [210,806,873, 211,237,618]
Cisgender Men	**47.26%** 102,107,460	[47.20%, 47.31%] [101,951,911, 102,263,008]
Cisgender Women	**50.41%** 108,914,786	[50.36%, 50.46%] [108,773,855, 109,055,718]
Gender Minority	**2.34%** 5,049,763	[2.30%, 2.38%] [4,962,965, 5,136,561]
Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. Gender minority refers to transgender and gender-expansive (something else) individuals.		

Table 6: Sexual and Gender Identity Estimates of U.S. Adults and Within-SGM Sexual and Gender Identity Estimates (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)		
	% Population Total	[95% CI]
SGM	**12.03%** 25,575,448	[11.94%, 12.12%] [25,391,564, 25,759,331]
(within-SGM)		
Straight	**4.83%**	[4.66%, 5.00%]
Sexual Minority	**95.17%**	[95.00%, 95.34%]
Lesbian or gay	**30.07%**	[29.66%, 30.48%]
Bisexual+	**65.11%**	[64.68%, 65.53%]
Gender Minority	**21.32%**	[21.00%, 21.65%]
Cisgender	**78.68%**	[78.35%, 79.00%]
Cisgender Men	**33.51%**	[33.20%, 33.82%]
Cisgender Women	**45.17%**	[44.79%, 45.55%]
Non-SGM	**87.97%** 187,036,744	[87.88%, 88.06%] [186,776,106, 187,297,382]

Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Population totals exclude respondents with "don't know" selected as their sexual identity. Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. Totals for within-SGM /non-SGM estimates excluded intentionally. Straight SGM people are transgender individuals and are intentionally included (for example, straight-identifying transgender men and women). Gender minority refers to transgender and gender-expansive (something else) individuals.

Table 7: Age Distribution Among U.S. Adults and within-Sexual and Gender Identity (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)			
	U.S. Adults	**Non-SGM**	**SGM**
	% [95% CI]	% [95% CI]	% [95% CI]
18-24	**7.48%** [7.39%, 7.56%]	**5.81%** [5.74%, 5.89%]	**19.64%** [19.27%, 20.02%]
25-34	**18.55%** [18.47%, 18.63%]	**16.41%** [16.33%, 16.50%]	**34.15%** [33.79%, 34.52%]
35-44	**18.17%** [18.11%, 18.24%]	**18.16%** [18.08%, 18.24%]	**18.27%** [18.06%, 18.49%]
45-54	**16.36%** [16.30%, 16.41%]	**17.18%** [17.12%, 17.25%]	**10.33%** [10.13%, 10.53%]
55-64	**17.12%** [17.07%, 17.16%]	**18.23%** [18.17%, 18.29%]	**8.97%** [8.78%, 9.17%]
65+	**22.33%** [22.27%, 22.39%]	**24.20%** [24.14%, 24.27%]	**8.63%** [8.44%, 8.82%]

Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. N = 2,148,860

Table 8: Educational Level Distribution Among U.S. Adults and within-Sexual and Gender Identity (U.S. Household Pulse Survey Pooled Estimate, 2021-2024) 12.			
	U.S. Adults	**Non-SGM**	**SGM**
	% [95% CI]	% [95% CI]	% [95% CI]
High school or less	**36.37%** [36.30%, 36.43%]	**36.81%** [36.73%, 36.88%]	**33.14%** [32.66%, 33.62%]
Associates or some college	**30.35%** [30.30%, 30.39%]	**29.77%** [29.71%, 29.82%]	**34.57%** [34.20%, 34.95%]
Bachelor	**18.36%** [18.32%, 18.40%]	**18.28%** [18.24%, 18.33%]	**18.91%** [18.69%, 19.14%]
Graduate	**14.93%** [14.88%, 14.98%]	**15.14%** [15.09%, 15.20%]	**13.37%** [13.19%, 13.56%]
Note: Based on Household Pulse Survey Data weeks 34-63 and cycles 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. N = 2,148,860			

Table 9: Regional Distribution Among U.S. Adults and within-Sexual and Gender Identity (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)			
	U.S. Adults	**Non-SGM**	**SGM**
	% [95% CI]	% [95% CI]	% [95% CI]
Northeast	**17.26%** [17.22%, 17.31%]	**17.26%** [17.20%, 17.33%]	**17.26%** [16.92%, 17.61%]
South	**39.00%** [38.95%, 39.05%]	**39.34%** [39.27%, 39.42]	**36.47%** [36.12%, 36.82]
Midwest	**21.20%** [21.17%, 21.23%]	**21.38%** [21.33%, 21.42%]	**19.88%** [19.66%, 20.10%]
West	**22.54%** [22.50%, 22.59%]	**22.01%** [21.95%, 22.08%]	**26.39%** [26.00%, 26.77%]
Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. N = 2,148,860			

Table 10: Racial and Ethnic Distribution Among U.S. Adults and within-Sexual and Gender Identity (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)			
	U.S. Adults	**Non-SGM**	**SGM**
	% [95% CI]	% [95% CI]	% [95% CI]
Asian	**5.39%** [5.33%, 5.42%]	**5.56%** [5.50%, 5.61%]	**4.08%** [3.91%, 4.26%]
Black	**11.74%** [11.70%, 11.79%]	**12.06%** [12.00%, 12.11%]	**9.47%** [9.24%, 9.71%]
Hispanic	**13.28%** [13.21%, 13.34%]	**12.64%** [12.56%, 12.72%]	**17.92%** [17.59%, 18.26%]
White	**65.32%**	**65.78%**	**61.95%**

	[65.26%, 65.38%]	[65.69%, 65.87%]	[61.52%, 62.38%]
Other	**4.28%** [4.24%, 4.32%]	**3.96%** [3.92%, 4.01%]	**6.58%** [6.37%, 6.79%]
Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files.			

Table 11: SGM Share of Population and Totals by State and DC (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)

	SGM % [95% CI]	SGM Population Total [95% CI]
Alabama	**9.49%** [9.00%, 10.01%]	**316,576** [299,708, 333,444]
Alaska	**14.13%** [13.52%, 14.76%]	**65,183** [62,238, 68,127]
Arizona	**13.34%** [12.92%, 13.76%]	**620,202** [600,673, 639,732]
Arkansas	**9.69%** [9.08%, 10.33%]	**191,564** [179,220, 203,907]
California	**14.01%** [13.57%, 14.45%]	**3,119,511** [3,021,311, 3,217,711]
Colorado	**14.38%** [13.92%, 14.86%]	**553,073** [534,418, 571,728]
Connecticut	**11.60%** [11.13%, 12.09%]	**272,561** [261,063, 284,058]
Delaware	**10.98%** [10.38%, 11.61%]	**74,949** [70,655, 79,243]
District of Columbia	**22.62%** [21.74%, 23.54%]	**102,773** [98,623, 106,922]
Florida	**11.39%** [10.98%, 11.82%]	**1,685,461** [1,622,288, 1,748,634]
Georgia	**11.46%** [10.99%, 11.96%]	**813,698** [778,836, 848,559]
Hawaii	**11.16%** [10.48%, 11.88%]	**105,322** [98,642, 112,001]
Idaho	**10.46%** [9.99%, 10.96%]	**129,835** [123,796, 135,874]
Illinois	**11.43%** [11.06%, 11.81%]	**911,441** [880,166, 942,717]
Indiana	**11.39%** [10.94%, 11.85%]	**505,641** [485,159, 526,122]
Iowa	**10.51%** [10.06%, 10.98%]	**219,443** [209,815, 229,070]
Kansas	**11.25%** [10.84%, 11.67%]	**211,430** [203,295, 219,566]
Kentucky	**10.95%** [10.40%, 11.51%]	**325,589** [309,418, 341,760]
Louisiana	**9.73%** [9.26%, 10.23%]	**288,185** [273,582, 302,788]
Maine	**12.60%** [12.01%, 13.22%]	**122,918** [116,991, 128,845]
Maryland	**11.70%** [11.30%, 12.11%]	**470,783** [453,902, 487,664]
Massachusetts	**13.91%** [13.53%, 14.30%]	**636,285** [618,329, 654,242]
Michigan	**10.95%** [10.59%, 11.33%]	**736,037** [710,236, 761,838]
Minnesota	**12.31%** [11.86%, 12.78%]	**461,438** [444,080, 478,796]
Mississippi	**7.98%** [7.35%, 8.67%]	**151,383** [138,592, 164,174]
Missouri	**11.55%** [11.12%, 12.00%]	**474,209** [456,036, 492,382]
Montana	**10.83%** [10.26%, 11.43%]	**82,862** [78,284, 87,440]
Nebraska	**11.16%** [10.75%, 11.58%]	**139,643** [134,457, 144,830]
Nevada	**13.56%** [13.04%, 14.11%]	**272,878** [261,777, 283,980]
New Hampshire	**11.88%** [11.31%, 12.48%]	**115,956** [110,111, 121,802]
New Jersey	**9.55%** [9.09%, 10.04%]	**554,769** [526,475, 583,063]

New Mexico	**12.97%** [12.34%, 13.63%]	**165,495** [157,091, 173,898]
New York	**13.08%** [12.50%, 13.68%]	**1,605,943** [1,530,639, 1,681,247]
North Carolina	**10.56%** [10.06%, 11.09%]	**747,819** [710,738, 784,900]
North Dakota	**9.97%** [9.31%, 10.66%]	**49,516** [46,181, 52,851]
Ohio	**11.29%** [10.87%, 11.72%]	**883,480** [849,750, 917,209]
Oklahoma	**11.58%** [11.05%, 12.12%]	**298,487** [284,802, 312,171]
Oregon	**17.29%** [16.85%, 17.74%]	**493,056** [480,117, 505,995]
Pennsylvania	**11.14%** [10.76%, 11.53%]	**956,487** [922,673, 990,301]
Rhode Island	**12.48%** [11.77%, 13.22%]	**88,860** [83,653, 94,067]
South Carolina	**9.82%** [9.26%, 10.42%]	**349,740** [328,953, 370,527]
South Dakota	**8.79%** [8.23%, 9.39%]	**51,579** [48,099, 55,059]
Tennessee	**10.65%** [10.24%, 11.07%]	**499,919** [480,429, 519,409]
Texas	**12.54%** [12.08%, 13.01%]	**2,243,590** [2,158,437, 2,328,742]
Utah	**12.88%** [12.45%, 13.33%]	**267,123** [257,840, 276,405]
Vermont	**13.76%** [13.12%, 14.43%]	**61,464** [58,484, 64,445]
Virginia	**11.38%** [10.90%, 11.87%]	**647,562** [619,867, 675,257]
Washington	**16.24%** [15.86%, 16.62%]	**836,099** [816,115, 856,083]
West Virginia	**9.79%** [9.13%, 10.49%]	**119,495** [111,080, 127,910]
Wisconsin	**11.15%** [10.75%, 11.57%]	**440,513** [424,559, 456,467]
Wyoming	**9.80%** [9.18%, 10.47%]	**37,623** [35,115, 40,131]

Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files.

Table 12: SGM Share of Population and Totals by Metropolitan Statistical Area, Largest 15 MSAs (U.S. Household Pulse Survey Pooled Estimate, 2021-2024)

Code	Name	SGM % [95% CI]	SGM Population Total [95% CI]
12060	Atlanta, Sandy Springs, Roswell	**12.63%** [12.02%, 13.26%]	**502,321** [476,640, 528,002]
14460	Boston, Cambridge, Newton	**14.20%** [13.70%, 14.73%]	**442,855** [426,356, 459,354]
16980	Chicago, Naperville, Elgin,	**12.07%** [11.60%, 12.55%]	**695,380** [667,642, 723,118]
19100	Dallas, Fort Worth, Arlington,	**13.11%** [12.47%, 13.78%]	**632,947** [599,148, 666,747]
19820	Detroit, Warren, Dearborn	**10.12%** [9.55%, 10.73%]	**297,718** [279,977, 315,460]
26420	Houston, The Woodlands, Sugar Land,	**12.77%** [12.03%, 13.55%]	**565,912** [528,823, 603,000]

31080	Los Angeles, Long Beach, Anaheim	**14.32%** [13.63%, 15.05%]	**1,069,253** [1,014,872, 1,123,635]
33100	Miami, Fort Lauderdale, West Palm Beach	**11.23%** [10.37%, 12.08%]	**452,651** [418,068, 487,234]
35620	New York, Newark, Jersey City	**12.27%** [11.78%, 12.78%]	**1,393,123** [1,333,923, 1,452,322]
37980	Philadelphia, Camden, Wilmington,	**11.40%** [10.96%, 11.85%]	**475,061** [455,066, 495,056]
38060	Phoenix, Mesa, Chandler	**13.72%** [13.19%, 14.28%]	**432,348** [415,002, 449,693]
40140	Riverside, San Bernardino, Ontario	**13.01%** [12.19%, 13.89%]	**371,976** [345,136, 398,815]
41860	San Francisco, Oakland, Berkeley	**17.08%** [16.22%, 17.97%]	**448,316** [422,198, 474,433]
42660	Seattle, Tacoma, Bellevue	**17.52%** [16.97%, 18.09%]	**472,763** [456,073, 489,452]
47900	Washington, Arlington, Alexandria	**12.81%** [12.35%, 13.29%]	**487,497** [468,981, 506,012]

Note: Based on Household Pulse Survey Data weeks 34-63 and phases 4.0-4.1 (2021-2024). Estimates are weighted. Confidence intervals are estimated using balanced repeated replication with Fay's coefficient of 0.5 and 80 replications. Replicate weights provided by the U.S. Census Bureau in public use files. Area names reflect the official U.S. Office of Management and Budget (OMB) titles for Core Based Statistical Areas and do not enumerate every city or county contained within each metropolitan area.

5. Supplemental Tables for Robustness Checks: LGBTQ+ Identification Within Demographic Groups

Table A. SGM and Non-SGM Share of Age Groups (U.S. Household Pulse Survey Pooled Estimate, 2021–2024)

	18–24	25–34	35–44	45–54	55–64	65+
Non-SGM	**68.39%**	**77.85%**	**87.90%**	**92.41%**	**93.69%**	**95.35%**
	[67.84%, 68.94%]	[77.61%, 78.09%]	[87.72%, 88.09%]	[92.25%, 92.56%]	[93.54%, 93.84%]	[95.25%, 95.45%]
SGM	**31.61%**	**22.15%**	**12.10%**	**7.593%**	**6.306%**	**4.648%**
	[31.06%, 32.16%]	[21.91%, 22.39%]	[11.91%, 12.28%]	[7.442%, 7.747%]	[6.160%, 6.456%]	[4.550%, 4.748%]

Table B. SGM and Non-SGM Share of Racial/Ethnic Groups (U.S. Household Pulse Survey Pooled Estimate, 2021–2024)					
	Asian	Black	Hispanic	White	Other
Non-SGM	**90.87%**	**90.30%**	**83.76%**	**88.59%**	**81.51%**
	[90.46%, 91.27%]	[90.05%, 90.54%]	[83.41%, 84.11%]	[88.50%, 88.69%]	[80.93%, 82.07%]
SGM	**9.13%**	**9.70%**	**16.24%**	**11.41%**	**18.49%**
	[8.729%, 9.542%]	[9.457%, 9.948%]	[15.89%, 16.59%]	[11.31%, 11.50%]	[17.93%, 19.07%]

Table C. SGM and Non-SGM Share of Education Level Group (U.S. Household Pulse Survey Pooled Estimate, 2021–2024)				
	High school or less	Associate or some college	Bachelor	Graduate
Non-SGM	**89.04%**	**86.30%**	**87.61%**	**89.23%**
	[88.83%, 89.24%]	[86.14%, 86.45%]	[87.47%, 87.74%]	[89.09%, 89.36%]
SGM	**10.96%**	**13.70%**	**12.39%**	**10.77%**
	[10.76%, 11.17%]	[13.55%, 13.86%]	[12.26%, 12.53%]	[10.64%, 10.91%]

Table D. SGM and Non-SGM Share of Regional Group (U.S. Household Pulse Survey Pooled Estimate, 2021–2024)				
	Northeast	South	Midwest	West
Non-SGM	**87.97%**	**88.75%**	**88.72%**	**85.92%**
	[87.70%, 88.24%]	[88.61%, 88.89%]	[88.59%, 88.84%]	[85.69%, 86.14%]
SGM	**12.03%**	**11.25%**	**11.28%**	**14.08%**
	[11.76%, 12.30%]	[11.11%, 11.39%]	[11.16%, 11.41%]	[13.86%, 14.31%]

6. Discussion

6.1. Selecting the Household Pulse Survey

No single federal or private survey provides a definitive count of the adult population by sexual and gender identity in the United States. Estimates vary across data sources because surveys differ in their target populations, sampling designs, modes of administration, question wording, response options, treatment of missing or uncertain responses, and periods of data collection. The objective of the present analysis was therefore not to identify one survey as providing the uniquely "true" population estimate. Rather, it was to identify the data source most suitable for producing stable, multidimensional population benchmarks for calibration of the Community Insights nonprobability panel, with totals as a byproduct of such estimation. The Household Pulse Survey was selected because it offered a combination of features not available together in the principal alternative surveys considered. The HPS was a national probability-based household survey administered online by the U.S. Census Bureau. Beginning in 2021, it included separate questions on sexual orientation, sex assigned at birth, and current gender identity. It also included detailed demographic variables and Census-provided person and replicate weights, allowing both weighted point estimation and design-consistent variance estimation. Most importantly for the present application, the repeated HPS cross-sections produced a very large cumulative sample. Pooling 37 administrations from 2021 through 2024 yielded 2,452,769 observations before exclusions for imputation and item nonresponse. The resulting sample supported the estimation of sexual identity, gender identity, and combined SOGI distributions within age, race and ethnicity, education, and region categories with substantially greater precision than would generally be possible from a single annual survey. Accordingly, the HPS was selected not because its estimates should automatically be interpreted as more accurate than estimates from other surveys, but because it was the most suitable available source for the specific purpose of developing stable calibration benchmarks. It uniquely combines a national probability design, self-administered data collection, separate SOGI measures, a large pooled sample, relevant demographic covariates, public-use microdata, and replicate weights for variance estimation.

6.2. Limitations of the Household Pulse Survey

The analysis, like any, is subject to limitations. First, the HPS was introduced as an experimental data product designed to rapidly measure emerging social and economic conditions. It was not originally designed specifically to enumerate the population by sexual and gender identity. Its rapid-response design, relatively short data-collection periods, and changing questionnaire structure distinguish it from longer-established federal surveys. Although the use of probability sampling, weighting, and replicate-weight variance estimation strengthens inference, these procedures cannot eliminate all potential coverage, nonresponse, or measurement error. Second, the HPS was administered online. Self-administration may provide greater privacy for sensitive questions, especially for sexual and gender minorities (McInroy, 2016), but exclusive reliance on online response can produce differential coverage or participation. The final person

weights adjust for known differences in selection and response, but weighting can correct only for characteristics represented in the adjustment process. Residual nonresponse bias may remain when survey participation is associated with SOGI identity after conditioning on the variables used in weighting. Third, the analysis pools cross-sections collected over four years. The resulting estimand is an average population distribution across the 37 included administrations rather than the population distribution at a single point in time. Pooling greatly improves precision but may obscure genuine change occurring between 2021 and 2024. Each cross-section was given equal influence through the common weight rescaling, regardless of whether the corresponding HPS administration covered an approximately two-week or monthly collection cycle. The estimates should therefore be understood as average cross-sectional benchmarks, not calendar-duration-weighted averages.

Fourth, questionnaire changes limited the period that could be harmonized. Phase 4.2 was excluded to maintain comparability with the SOGI measures used in earlier administrations. The resulting estimates apply to the retained Weeks 34–63 and Phase 4.0–4.1 Cycles 1–7 and should not be assumed to describe later HPS measurement without further harmonization. Fifth, the available gender-identity measures constrain classification. Although HPS employed separate sex-assigned-at-birth and current-gender questions, the current-gender item did not allow the analysis to identify all transgender and nonbinary subgroups with equal conceptual precision. Respondents selecting transgender or "none of these" could not always be reliably differentiated into transgender men, transgender women, nonbinary adults, or other gender-expansive identities. The broad gender-minority benchmark used in this analysis therefore prioritizes stable and defensible calibration categories over detailed representation of distinct identities. Sixth, estimates depend on the treatment of missing, imputed, and uncertain responses. Cases containing imputed values for sex assigned at birth, birth year, race and ethnicity, or education were excluded so that benchmark margins would be based on observed respondent reports. "I don't know" responses to the sexual-orientation item were excluded from the primary estimates and examined separately in sensitivity analyses. These decisions improve the interpretability of the resulting calibration controls but also change the target population represented by each estimate. If respondents with missing, imputed, or uncertain answers differ systematically from retained respondents, the estimates may be affected by item-nonresponse bias.

Seventh, very narrow confidence intervals should not be interpreted as evidence that total error is negligible. The very large, pooled sample produces high sampling precision, and the BRR estimates account for sampling variation represented by the replicate weights. They do not fully quantify nonsampling errors arising from coverage, nonresponse, question interpretation, identity disclosure, category construction, or differences between the HPS target population and the population represented in Community Insights. The reported confidence intervals therefore describe estimated sampling uncertainty under the HPS design, not the full range of uncertainty surrounding the underlying LGBTQ+ population quantities. Finally, calibration to HPS benchmarks cannot eliminate all selection bias in the Community Insights panel. This is perhaps most important to remember as survey researchers, no matter how well designed or meticulously thought through benchmarks are constructed. Post-stratification can align the panel with selected population margins, but it cannot correct differences between respondents and nonrespondents on characteristics that are unmeasured or omitted from the calibration model. The benchmarks should thus be viewed as an important component of a broader strategy for improving comparability and

reducing selection bias, rather than as a complete solution to the inferential limitations of nonprobability sampling.

6.3. Other Potential Data Sources Considered

6.3.1. Gallup Poll Social Series

Gallup provides one of the most prominent and longest-running sources of national estimates of LGBTQ+ identification. Its estimates are valuable for tracking broad changes in self-identification over time and are widely recognized in public discourse. Gallup's recent estimates are based on aggregated national telephone interviews. Its questionnaire asks respondents whether they identify as heterosexual, lesbian, gay, bisexual, transgender, or something else, allowing respondents to report one or more LGBTQ+ identities. Gallup was not selected as the primary benchmark source for several reasons. First, its telephone mode differs from the self-administered web mode used by HPS. Reporting of socially sensitive characteristics can vary by mode because respondents may experience different levels of privacy and interviewer presence. This does not establish that the HPS estimate is necessarily more accurate, but it provides one plausible explanation for differences between the surveys. Second, Gallup's combined identity question is not directly equivalent to the separate sexual and gender identity measures used in HPS. A respondent's answer may therefore reflect both substantive identity and the structure of the response task. Third, the annual Gallup samples used for recent LGBTQ+ estimates are much smaller than the pooled HPS sample. For example, Gallup's 9.0% estimate based on 2025 data drew on telephone interviews with more than 13,000 adults (Jones, 2026). Although this is adequate for estimating an overall national percentage, it provides less capacity for constructing the detailed cross-classified demographic margins needed for post-stratification. Gallup therefore serves as an important external statistic for the overall prevalence estimate, but it is less well suited to producing the extensive public-use calibration controls required for the present analysis.

6.3.2. Survey of Household Economics and Decisionmaking

The Federal Reserve Board's Survey of Household Economics and Decisionmaking (SHED) is an annual survey designed primarily to assess the financial and economic well-being of U.S. households. It includes public-use microdata and has supported research examining differences in economic well-being by sexual orientation (Carpenter et al, 2025). SHED provides a useful robustness comparison because it is administered through a self-administered survey environment and includes relevant demographic and SOGI information. In the present analysis, a quick robustness check was done with pooled SHED data from 2021 through 2023 and produced an estimated higher than Gallup but smaller than HPS. This difference should not be interpreted by itself as evidence that any of these estimates are erroneous. The two surveys differ in their substantive purpose, sampling and weighting procedures, exact measures, response options, survey context, and years of inclusion. SHED was not selected as the primary source because its annual

samples are substantially smaller than the pooled HPS dataset and provide fewer observations for estimating stable joint distributions across LGBTQ+ status and multiple demographic characteristics. Its primary design objective is also the study of household financial well-being rather than population surveillance or the production of detailed demographic calibration margins. Nevertheless, its broadly similar estimate provides evidence that the HPS result is within the range produced by another high-quality national survey using a self-administered mode.

6.3.3. The General Social Survey

The General Social Survey is a nationally representative probability survey administered by NORC and has measured social characteristics, attitudes, and behavior in the United States since 1972. It is especially valuable for long-term trend analysis, questionnaire continuity, and the study of relationships between demographic characteristics and social attitudes. The GSS has included sexual identity measures and has experimented with two-step measurement of sex assigned at birth and current gender identity (Lagos & Compton, 2021; Pao et al., 2025; Smith and Son, 2019). The GSS has several methodological strengths, including its probability sampling framework, extensive documentation, and strong emphasis on measurement quality. However, its sample size is not designed to support highly detailed calibration benchmarks for relatively small population groups. The number of LGBTQ+ respondents available within a single GSS cross-section would generally be insufficient to produce precise intersections of SOGI status with age, race and ethnicity, education, and geography. Pooling multiple GSS years could improve precision, but changes in data-collection mode during and after the COVID-19 pandemic would require careful harmonization and interpretation. NORC itself cautions that observed changes in recent GSS waves may reflect both substantive change and methodological changes in survey administration. The GSS is therefore highly valuable for substantive and longitudinal social research but is less appropriate than HPS for constructing the high-precision, multidimensional population controls required here.

6.3.4. The Behavioral Risk Factor Surveillance System

The Behavioral Risk Factor Surveillance System is a large telephone-based health surveillance system coordinated by the Centers for Disease Control and Prevention. It collects data in all 50 states, the District of Columbia, and participating territories and is designed primarily to produce state-level estimates of health behaviors, chronic conditions, health-care access, and preventive-service use. BRFSS has major advantages for LGBTQ+ population research. Its cumulative sample is large, its state-based design permits geographic analysis that is not possible in many national surveys, and its SOGI measures have been used extensively to estimate health and demographic outcomes among sexual and gender minority adults. It is therefore an especially strong source when the analytic objective is state-level prevalence estimation or examination of health disparities. However, SOGI questions have historically been administered through optional modules rather than uniformly by every jurisdiction every year (Jesdale, 2021). A national estimate constructed from BRFSS consequently requires decisions about which states and years to include, how to address variation in module participation, and whether the included jurisdictions

adequately represent the national population. Its telephone mode also differs from the web-based HPS mode. BRFSS remains a valuable alternative for validation and sensitivity analysis, particularly for state-level or health-focused analyses. For the present purpose, however, HPS offered a more straightforward nationally weighted source with SOGI measures administered within a common federal survey framework and a larger harmonized sample for estimating national calibration margins. Nevertheless, BRFSS data paired with robust statistical techniques, like small area estimation, have been employed for providing figures of sexual and gender identity prevalence among adults in the U.S. previously (see Flores & Conron, 2023).

6.3.5. American Community Survey

The American Community Survey (ACS) represented perhaps the most promising prospective source for benchmarking national LGBTQ+ population estimates. As the nation's largest continuous household survey, the ACS offers unparalleled sample sizes, comprehensive geographic coverage, and a well-established weighting framework that supports reliable estimates for states, metropolitan areas, and many smaller geographic areas. In 2024, the U.S. Census Bureau conducted the American Community Survey (ACS) Sexual Orientation and Gender Identity (SOGI) Methods Panel Test (United States Census Bureau, 2024) to evaluate the feasibility of incorporating SOGI measures into the nation's largest annual household survey. The large-scale methodological study experimentally evaluated alternative question wording, confirmation questions, write-in response formats, and measurement quality across survey modes, self- and proxy-response, and demographic subgroups. The Bureau indicated that the results would inform potential implementation of SOGI questions in the production ACS and that methodological findings would be publicly released. However, following the 2025 change in presidential administration, implementation of SOGI measures in the ACS did not proceed, and the anticipated evaluation report has not been publicly released. Consequently, the ACS did not become the comprehensive federal benchmark for LGBTQ+ population estimates that many researchers had anticipated.

7. Summary

Despite these limitations, the HPS provides the strongest available foundation for the present calibration objective. Gallup offers an important and recognizable indicator of national LGBTQ+ identification; SHED provides a valuable robustness comparison using a self-administered national survey; the GSS offers strong probability sampling and high-quality social measurement; and BRFSS offers exceptional cumulative sample size and state-level surveillance capacity. None, however, combines the HPS's pooled sample size, separate SOGI measures, national weighting framework, demographic detail, public-use respondent data, and replicate weights in a form equally suited to constructing the required calibration benchmarks. The HPS estimates should therefore be treated neither as a definitive census of LGBTQ+ adults nor as automatically superior to estimates from other surveys. But they are the estimates of the Human Rights Campaign Foundation and best understood as methodologically transparent and highly precise, derived from

the data source assessed to be most fit for the intended purpose: constructing stable national SOGI benchmarks for calibration of a large LGBTQ+ nonprobability research panel.

8. Authors

Charleigh J. Flohr
Associate Director, Research, Human Rights Campaign Foundation

Tari Hanneman
Senior Director, Research & Education, Human Rights Campaign Foundation

9. References

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology, 1*(2), 90–143. https://doi.org/10.1093/jssam/smt008

Carpenter, C. S., Dasgupta, K., Merchant, Z., & Plum, A. (2025). Sexual orientation and financial well-being in the United States. *Journal of Economic Behavior & Organization, 233*, 106967. https://doi.org/10.1016/j.jebo.2025.106967

Dever, J. A., Rafferty, A., & Valliant, R. (2008). Internet surveys: Can statistical adjustments eliminate coverage bias? *Survey Research Methods, 2*(2), 47–62.

Deville, J.-C., & Sarndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87*(418), 376–382.

Flores, A. R., & Conron, K. J. (2023). *Adult LGBT population in the United States*. Williams Institute, UCLA School of Law.

Gallup. (2025, February 20). *LGBTQ+ identification holds at 9.3% in U.S.* https://news.gallup.com/poll/702206/lgbtq-identification-holds.aspx

Jesdale, B. M., Gilbert, L. K., Wood, S., McGranahan, M., Hughes, A., Pizacani, B. A., & Fish, J. N. (2021). Sources of missing sexual orientation and gender identity data in the Behavioral Risk Factor Surveillance System. *American Journal of Preventive Medicine, 61*(2), 281–290. https://doi.org/10.1016/j.amepre.2021.02.027

Julian, C. A., Manning, W. D., & Westrick-Payne, K. K. (2024). Responses to sexual and gender identity measures in population-level data by birth cohort: A research note. *Demography, 61*(1), 15–30. https://doi.org/10.1215/00703370-11164985

Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics, 19*(2), 81–97.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Lagos, D., & Compton, D. (2021). Evaluating the use of a two-step gender identity measure in the 2018 General Social Survey. *Demography, 58*(2), 763–772. https://doi.org/10.1215/00703370-8976151

McInroy, L. B. (2016). Pitfalls, potentials, and ethics of online survey research: LGBTQ and other marginalized and hard-to-access youths. *Social Work Research, 40*(2), 83–94. https://doi.org/10.1093/swr/svw005

Pao, C., Julian, C. A., Compton, D., Lagos, D., & Stacey, L. (2025). Demographic differences in responses to a two-step gender identity measure. *Sociological Science, 12*, 277–293. https://doi.org/10.15195/v12.a13

Smith, T. W., & Son, J. (2019). *Transgender and alternative gender measurement on the 2018 General Social Survey* (GSS Methodology Report No. 129). NORC at the University of Chicago.

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. Oxford University Press.

United States Census Bureau (2022). Release Note on Using the Household Pulse Survey to Estimate Transgender Respondents. *Phase 3.3 Household Pulse Survey User Notes.* U.S. Department of Commerce: Phase3-3_2021_Household_Pulse_Survey_User_Notes_02162022.pdf

United States Census Bureau. (2024). *Supporting Statement B: 2024 American Community Survey (ACS) Sexual Orientation and Gender Identity (SOGI) Test*. U.S. Department of Commerce. 2024-09101.pdf