Center for the Study of Democratic Institutions (CSDI)

Vanderbilt Poll Spring 2024
Methods Report

ssrs
redefining research

Table of Contents

Summary

The Center for the Study of Democratic Institutions (CSDI) at Vanderbilt University regularly conducts public opinion polls of Tennessee registered voters to provide non-partisan, scientifically based public opinion data. Each year, the Vanderbilt Poll conducts at least two statewide surveys, one prior to the start of the legislative season and one at its conclusion. These polls provide point-in-time data to find out what registered voters in Tennessee think about national, state, and local public policy issues.

The Tennessee Poll Spring 2024 obtained telephone interviews with a representative sample of 1,003 registered voters, ages 18 or older, living in Tennessee. Telephone interviews were conducted by landline (402) and cell phone (601, including 523 without a landline phone). Interviews were done in English from April 26-May 9, 2024. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 3.4 percentage points.

Details on the design, execution and analysis of the survey are discussed below.

Questionnaire Development

The questionnaire was developed by the Principal Investigators at the Center for the Study of Democratic Institutions (CSDI) at Vanderbilt in consultation with the SSRS project team. Prior to the field period, SSRS programmed the study into Forsta Plus (formerly known as Confirmit) Computer Assisted Telephone Interviewing (CATI) software. Extensive checking of the program was conducted to ensure that skip patterns and sample splits followed the design of the questionnaire.

Sample Design

The target population for this poll was Tennessee registered voters ages 18 or older. SSRS used a registration-based sampling (RBS) approach. RBS samples were procured from Aristotle, one of the major providers of voter list samples. Samples were provided according to SSRS specifications.

The sample frame was split into four strata based on age: [1] 18-29; [2] 30-44; [3] 45-64; [4] 65+. Samples were drawn within stratum regardless of whether they had a phone number appended. Records that had no phone number were sent to Dynata to get cell phone and landline numbers appended. The following table shows how much sample was released in each stratum after the Dynata phone append.

Table 1: Sample Released

Stratum	Landline	Cell	Total
18-29	12,212	8,704	20,916
30-44	10,979	14,432	25,411
45-64	13,938	9,145	23,083
65+	11,024	2,566	13,590
Total	48,153	34,847	83,000

The sample size for this poll was n=1,003 interviews. Overall, 40% of respondents were reached via landline and 60% of respondents were reached via cellular telephone.

Contact Procedures and Data Processing

Contact Procedures

For respondent selection for landline sample, interviews were conducted with the youngest adult male/female, ages 18 or older, who was at home based on a random rotation. For the cellular sample, interviews were conducted with the person who answered the phone. Interviewers verified that the person was an adult and in a safe place before administering the survey. For both landline and cell samples, after an adult was on the phone, geographic eligibility and current voter registration status were determined prior to accepting the respondent into the survey.

Survey Administration

The field period for this study was April 26-May 9, 2024. All interviews were completed in English using the CATI system. The CATI system ensured that questions followed logical skip patterns and that complete dispositions of all call attempts were recorded.

CATI interviewers received written materials about the survey instrument and received formal training for this particular project. The written materials were provided prior to commencement of data collection and included an annotated questionnaire that contained information about the goals of the study, detailed explanations about why questions were being asked, the meaning and pronunciation of key terms or names, potential obstacles to overcome in getting good answers to questions, and respondent problems that could be anticipated ahead of time, as well as strategies for addressing the potential problems.

Interviewer training was conducted before the study was launched. Interviewers were given instructions to help them maximize response rates and ensure accurate data collection.

To maximize survey response, SSRS enacted the following procedures during the field period:

- As many as five (5) attempts were made to contact every sampled telephone number.
- Calls were staggered over times of day and days of the week to maximize the chance of making contact with potential respondents. At least one daytime call was conducted if necessary.
- Interviewers explained the purpose of the study and its importance.
- Respondents were offered the option of scheduling a call-back at their convenience.
- Specially trained interviewers contacted numbers where the initial call resulted in respondents hanging up the phone.

Data Processing and Integration

Prior to running cross-tabulations, data were cleaned and checked using standard procedures. This program establishes editing parameters to locate any errors. Minimal back-coding was conducted for Race and self-reported county. No other coding was done for open-end responses.

Weighting And Analysis

Data were weighted to represent registered voters in Tennessee. The data were weighted by applying a base weight and balancing the demographic profile of the sample to target population parameters.

Base Weights

The sample frame was divided into 16 strata based on the region (East, Nashville Area, Central, Memphis/West) and age (18-29, 30-44, 45-64, 65+) of each sample record. The sample was disproportionately allocated across strata to try and compensate for the lower response typically seen among younger voters while also controlling the regional distribution of the sample.

The base weight of each respondent i is $BW_i = P_i/p_i$ where P_i is the proportion of the sample frame in stratum i and p_i is the proportion of interviews conducted in stratum i.

Calibration

With the base weight applied, the data were weighted to balance the demographic profile of the sample to the target population parameters.

Missing data in the raking variables were imputed using hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. Hot decking was done using an SPSS macro detailed in 'Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handing Missing Data' (Myers, 2011).

Weighting was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure.[1]

Data were weighted to distributions of sex, age, race, education, and region. The following table shows the data sources used for calibration totals.

Table 2: Calibration Variable Sources

Dimensions	Source
Sex	
Age	November 2022 Current Population Survey Voting and
Education	Registration Supplement[2]
Race	
Region	2018-2022 American Community Survey 5-Year Estimates[3]

Weights were trimmed at the 2nd and 98th percentiles to prevent individual interviews from having too much influence on survey-derived estimates. The table below compares unweighted and weighted sample distributions to target population benchmarks.

---

[1] https://community.ibm.com/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=17fd2f0b-7555-6ccd-c00c-5388b082161b&forceDialog=0
[2] Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler and Michael Westberry. IPUMS CPS: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2023. https://doi.org/10.18128/D030.V11.0
[3] https://www.census.gov/topics/research/guidance/planning-databases.html

Table 3: Sample Demographics

CATEGORY	VALUES	PARAMETER	UNWEIGHTED	WEIGHTED
SEX	Male	47.7%	46.5%	47.5%
	Female	52.3%	53.5%	52.5%
EDUCATION	18-29	13.1%	8.0%	12.6%
	30-34	6.8%	4.0%	6.6%
	35-44	17.6%	16.2%	17.7%
	45-54	16.7%	14.1%	16.8%
	55-64	17.1%	19.2%	17.3%
	65+	28.8%	38.6%	29.1%
EDUCATION	HS grad or less	32.1%	29.4%	32.2%
	Some college	30.2%	20.4%	29.7%
	College+	37.7%	50.1%	38.1%
RACE/ETHNICITY	White, Non-Hisp	80.1%	83.4%	80.4%
	Other	19.9%	16.6%	19.6%
REGION	East	36.6%	36.9%	36.5%
	Nashville Area	24.8%	24.7%	24.8%
	Central	20.8%	20.6%	20.8%
	Memphis/West	17.8%	17.7%	18.0%

Effects of Sample Design on Statistical Inference

Post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. SSRS calculates the effects of these design features so that an appropriate adjustment can be incorporated into tests of statistical significance when using these data. The so-called "design effect" or *deff* represents the loss in statistical efficiency that results from a disproportionate sample design and systematic non-response. The total sample design effect for this survey is 1.21.

SSRS calculates the composite design effect for a sample of size n, with each case having a weight, w, as:[4]

$$deff = \frac{n \sum w^2}{(\sum w)^2}$$

The survey's margin of error is the largest 95% confidence interval for any estimated proportion based on the total sample — the one around 50%. For example, the margin of error for the entire

---

[4] Kish, L. (1992). Weighting for Unequal Pi. Journal of Official Statistics, Vol. 8, No.2, 1992, pp. 183-200.

sample is ± 3.4 percentage points. This means that in 95 out of every 100 samples drawn using the same methodology, estimated proportions based on the entire sample will be no more than 3.4 percentage points away from their true values in the population. Margins of error for subgroups will be larger. It is important to remember that sampling fluctuations are only one possible source of error in a survey estimate. Other sources, such as respondent selection bias, questionnaire wording, and reporting inaccuracy, may contribute additional error of greater or lesser magnitude.

Response Rates

Table 4 reports the disposition of all sampled telephone numbers ever dialed from the original telephone number samples. The response rate estimates the fraction of all eligible sample that was ultimately interviewed. Response rates are computed according to American Association for Public Opinion Research standards.[5]

- The RR3 for RBS landline was 7.7 percent.
- The RR3 for RBS cell was 3.6 percent.

---

[5] The American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR.

Table 4: Sample Disposition

	LAND-LINE	CELL
Eligible, Interview (Category 1)		
Complete	402	601
Eligible, Non-interview (Category 2)		
Refusal and breakoff	228	239
Refusal	0	0
Respondent never available	0	0
Telephone answering device (confirming HH)	75	111
Answering machine household-no message left	0	0
Answering machine household-message left	0	0
Other, non-refusals	0	0
Deceased respondent	0	0
Physically or mentally unable/incompetent	5	15
Language problem	5	99
Miscellaneous	0	0
Unknown eligibility, non-interview (Category 3)		
Busy/No answer	6,457	6,532
Answering machine-don't know if household	7,941	14,680
Call blocking	49	244
Technical phone problems	5,896	0
Residential, unknown if eligible respondent	0	0
No screener completed, residential and live contact made	2,419	7,179
No screener completed, residential and no live contact	0	0
Other	0	0
Not eligible (Category 4)		
Fax/data line	185	147
Non-working number	24,361	4,428
Nonresidence	89	212
Business, government office, other organizations	0	0
No eligible respondent	39	360
Quota filled	0	0
Other	0	0
Summary Dispositions		
I=Complete Interviews (1.1)	402	601
P=Partial Interviews (1.2)	0	0
R=Refusal and break off with eligible case (2.1)	228	239
NC=Non-contact with eligible case (2.2)	75	111
O=Other non-interview with eligible case (2.0, 2.3)	10	114
UH=Unknown if residential (3.0, 3.1)	20,345	21,456

UO=Unknown other (3.2, 3.9) (residential, unknown if eligible)	2,419	7,179
INNR = Ineligible: Not residential (4.0,4.1,4.2,4.3,4.4,4.5,4.8,4.9)	24,635	4,787
INR=Ineligible: Residential but ineligible for survey (4.7)	39	360
Total	48,153	34,847
ADDRESSING CASES WITH UNDETERMINED ELIGIBILITY		
e1 = the % of known-residential cases estimated to have eligible R	94.8%	74.7%
e2 = the % of unknown-if-residential cases that are estimated to be residential	11.4%	64.3%
Total sample used	**48,153**	**34,847**
Response Rate 3 (I / (I+P+R+NC+O+[e1*e2*UH]+[e1*UO]))	**7.7%**	**3.6%**
Cooperation Rate 3 ((I+INR)/(I+INR+R+(e2*UO)))	**14.3%**	**11.5%**
Refusal Rate 3 (R/(I+P+R+NC+O))	**84.5%**	**90.0%**
Contact Rate 3 ((I+P+R+O) / (I+P+R+O+NC))	**20.4%**	**11.6%**

Deliverables

SSRS delivered to Vanderbilt University:

- Final questionnaire instrument.
- Weighted dataset in SPSS.
- Weighted banners in PDF.
- Topline.
- A detailed methods report.

About SSRS

SSRS is a full-service survey and market research firm managed by a core of dedicated professionals with advanced degrees in the social sciences. Service offerings include the SSRS Opinion Panel and other Online Solutions, SSRS Business Insights, SSRS Virtual Insights, and SSRS Text Message Panel, as well as custom research programs – all driven by a central commitment to methodological rigor. The SSRS team is renowned for its multimodal approach, as well as its sophisticated and proprietary sample designs. Typical projects for the company include complex strategic, tactical, and public opinion initiatives in the U.S. and in more than 40 countries worldwide. Visit www.ssrs.com for more information.