

Methodology

Tennessee Poll Fall 2017

Prepared by SSRS for Vanderbilt University

December 2017

SUMMARY

The Tennessee Poll Fall 2017 obtained telephone interviews with a representative sample of 1,013 registered voters in Tennessee. Telephone interviews were conducted by landline (813) and cell phone (200). The survey was conducted by SSRS of Glen Mills, PA. Interviews were administered in English from November 16 to December 5, 2017. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 3.7 percentage points.

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

DESIGN AND DATA COLLECTION PROCEDURES

Questionnaire Design

The questionnaire was developed by Vanderbilt in consultation with the SSRS project team. Prior to the field period, SSRS programmed the study into CfMC 8.6 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

Samples were provided by Marketing Systems Group (MSG) according to SSRS specifications. Samples were pulled from the Tennessee State voter file. The database includes approximately 3.6 million records in total and approximately 917,000 with a phone number. The sample frame was split into twelve strata: [1] 18-29 with landline; [2] 30-34 with landline; [3] 35-44 with landline; [4] 45-54 with landline; [5] 55-64 with landline; [6] 65+ with landline; [7] 18-29 with only cell phone; [8] 30-34 with only cell phone; [9] 35-44 with only cell phone; [10] 45-54 with only cell phone; [11] 55-64 with only cell phone; and [12] 65+ with

only cell phone. Records with both a landline and cell phone were included in the landline strata. A separate sample was drawn in each stratum.

Survey Administration

CATI interviewers received written materials about the survey instrument and received formal training for this particular project. The written materials were provided prior to the beginning of the field period and included an annotated questionnaire that contained information about the goals of the study, detailed explanations as to why questions were being asked, the meaning and pronunciation of key terms, responses to commonly asked questions, potential obstacles to be 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.

Contact Procedures

Interviews were conducted from November 16-December 5, 2017. All interviews were completed through the CATI system. The CATI system ensured that questions followed logical skip patterns and that complete dispositions of all call attempts were recorded.

As many as five attempts were made to contact every sampled telephone number. Sample was released for interviewing in replicates, which are representative subsamples of the larger sample. Using replicates to control the release of sample ensures that complete call procedures are followed for the entire sample. Calls were staggered over times of day and days of the week to maximize the chance of making contact with potential respondents. Each phone number received at least one daytime call when necessary.

For the landline sample, interviewers asked to speak with the person named in the sample file. If there were two or more respondents with the same name at that number, interviewers asked for the respondent who is registered to vote at that address. For the cellular sample, interviews were conducted with the person named in the sample file. Interviewers verified that the person was in a safe place before administering the survey.

Once the target respondent was on the phone, interviewers confirmed that they lived in Tennessee before conducting the full interview.

WEIGHTING AND ANALYSIS

Weighting is generally used in survey analysis to compensate for sample designs and patterns of non-response that might bias results. The weighting ensures that the demographic profile of the sample

matches the profile of the target population. The data was weighted to match Tennessee registered voter parameters.

The first step of the weighting was to compute a base weight to account for the disproportionate sampling across the different strata. Base weights were computed so that the base weighted distribution of the sample drawn across strata is the same as the distribution of the entire sample frame across the strata.

The second stage of weighting balanced sample demographics to population parameters. The sample was balanced to match Tennessee registered voter parameters for sex, age, race, education, and region¹. The basic weighting parameters came from a special analysis of the Current Population Survey, November 2016: Voting and Registration Supplement. The region parameter came from the U.S. Census Bureau's 2011-2015 5-Year American Community Survey (ACS) data.

Weighting was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure. Weights were trimmed to prevent individual interviews from having too much influence on the final results. The use of these weights in statistical analysis ensures that the demographic characteristics of the sample closely approximate the demographic characteristics of the population. Table 1 compares weighted and unweighted sample distributions to population parameters.

¹ Tennessee counties were divided into four regions. The regions were East, Nashville area, Central, and Memphis/West.

Table 1: Sample Demographics

	<u>Parameter</u>	<u>Unweighted</u>	<u>Weighted</u>
<u>Gender</u>			
Male	46.3	45.7	46.1
Female	53.7	54.3	53.9
<u>Age</u>			
18-29	14.7	5.1	12.6
30-34	8.5	3.6	8.0
35-44	15.8	9.4	16.0
45-54	16.2	16.5	16.8
55-64	20.3	20.3	21.1
65+	24.5	45.1	25.4
<u>Race/Ethnicity</u>			
White/not Hispanic	80.0	85.8	80.2
Other	20.0	14.2	19.8
<u>Education</u>			
HS Grad or less	41.6	30.8	40.1
Some College/Assoc Degree	24.2	23.7	24.4
College Graduate	34.2	45.5	35.5
<u>Region</u>			
East	36.9	38.1	37.0
Nashville	23.5	24.3	23.5
Central	20.4	23.0	20.6
Memphis/West	19.2	14.6	19.0

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 1.43.

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

$$deff = \frac{n \sum_{i=1}^n w_i^2}{\left(\sum_{i=1}^n w_i \right)^2} \quad \text{formula 1}$$

In a wide range of situations, the adjusted *standard error* of a statistic should be calculated by multiplying the usual formula by the square root of the design effect (\sqrt{deff}). Thus, the formula for computing the 95% confidence interval around a percentage is:

$$\hat{p} \pm \left(\sqrt{deff} \times 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right) \quad \text{formula 2}$$

where \hat{p} is the sample estimate and n is the unweighted number of sample cases in the group being considered.

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 sample is ± 3.7 percentage points. This means that in 95 out every 100 samples drawn using the same methodology, estimated proportions based on the entire sample will be no more than 3.7 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 RATE

Table 2 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.² Thus the response rate for the landline samples was 6 percent. The response rate for the cellular samples was 5 percent.

² The American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR.

Table 2. Sample Disposition

<u>Landline</u>	<u>Cell</u>	
116	75	Non-residential/Business
521	690	Listed person does not exist/deceased
4	--	Cell in landline frame
6	1	Duplicate
647	766	OF = Out of Frame
2,873	594	Not working
240	2	Computer/fax/modem
3,113	596	NWC = Not working/computer
6,078	639	UHU _{NC} = Non-contact, unknown if household/unknown other
3,585	3,037	Voice mail
426	114	Privacy Manager
4,011	3,151	UO _{NC} = Non-contact, unknown eligibility
1,899	591	Refusals
2,208	620	Callbacks
112	52	Not available for duration of study
4,219	1,263	UO _R = Refusal, unknown if eligible
33	23	O = Other (language)
141	103	SO = Screen out (ineligible)
508	105	R = Refusal, known eligible (breakoffs and qualified CBs)
813	200	I = Completed interviews
19,563	6,846	T = Total numbers sampled

Table 2. Sample Disposition (continued)

72.1%	78.1%	$e1 = (I+R+SO+O+UO_R+UO_{NC})/(I+R+SO+O+UO_R+UO_{NC}+OF+NWC)$ - Est. frame eligibility of non-contacts
90.4%	74.8%	$e2 = (I+R)/(I+R+SO)$ - Est. screening eligibility of unscreened contacts
40.5%	31.7%	$CON = [I + R + (e2*[O + UO_R])]/[I + R + (e2*[O + UO_R + UO_{NC}] + (e1*e2*UHUO_{NC})]$
15.7%	15.8%	$COOP = I/[I + R + (e2*[O + UO_R])]$
6.4%	5.0%	$AAPOR\ RR3 = I/[I+R+(e2*(UO_R+UO_{NC}+O))]+[e1*e2*UHUO_{NC}] = CON*COOP$