Methodology

Tennessee Poll Spring 2017

Prepared by Princeton Survey Research Associates International
for Vanderbilt University

May 2017

SUMMARY

The Tennessee Poll Spring 2017 obtained telephone interviews with a representative sample of 1,004 registered voters in Tennessee. Telephone interviews were conducted by landline (811) and cell phone (193). The survey was conducted by Princeton Survey Research Associates International (PSRAI). Interviews were done in English by Princeton Data Source (PDS) and Survey Sampling International (SSI) under the direction of PDS from May 4 to 15, 2017. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 3.3 percentage points.

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

DESIGN AND DATA COLLECTION PROCEDURES

Sample Design

Samples were provided by Marketing Systems Group (MSG) according to PSRAI 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 eight strata: [1] 18-29 with landline; [2] 30-44 with landline; [3] 45-64 with landline; [4] 65+ with landline; and [5] 18-29 with only cell phone; [6] 30-44 with only cell phone; [7] 45-64 with only cell phone; [8] 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.

Contact Procedures

Interviews were conducted from May 4 to 15, 2017. 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 weighting balanced sample demographics to population parameters. The sample was balanced, by form, to match Tennessee registered voter parameters for sex, age, race, education, and region[1]. 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

---

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

population. Table 1 compares weighted and unweighted sample distributions to population parameters.

Table 1: Sample Demographics

	Parameter	Unweighted	Weighted
Gender			
Male	46.3%	47.5%	46.1%
Female	53.7%	52.5%	53.9%
Age			
18-29	14.7%	12.3%	15.1%
30-34	8.5%	3.3%	5.8%
35-44	15.8%	14.2%	16.2%
45-54	16.2%	14.2%	16.3%
55-64	20.3%	24.8%	21.0%
65+	24.5%	31.2%	25.6%
Race/Ethnicity			
White/not Hispanic	80.0%	81.8%	80.1%
Other	20.0%	18.2%	19.9%
Education			
HS Grad or less	41.6%	31.9%	40.5%
Some College/Assoc Degree	24.2%	25.9%	24.2%
College Graduate	34.2%	42.2%	35.2%
Region			
East	36.9%	37.6%	37.1%
Nashville	23.5%	21.5%	22.9%
Central	20.4%	24.4%	20.8%
Memphis/West	19.2%	16.4%	19.1%

Effects of Sample Design on Statistical Inference

Post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. PSRAI 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 1.16.

PSRAI calculates the composite design effect for a sample of size *n*, with each case having a weight, w_i as:

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

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:

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

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 ±3.3 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 3.3 percentage points away from their true values in the population. Margins of error for subgroups will be larger. The margin of error for results based on form 1 respondents is ±4.6 percentage points and the margin of error for form 2 respondents is ±4.9 percentage points. 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 11 percent. The response rate for the cellular samples was 7 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

Landline	Cell	
243	68	Non-residential/Business
48	11	Over quota
1,124	990	Listed person does not exist/deceased
1	----	Cell in landline frame
1,416	1,069	OF = Out of Frame
2,625	530	Not working
319	2	Computer/fax/modem
2,944	532	NWC = Not working/computer
892	25	NA/Busy all attempts
0	837	VM not set up/caller out of range
892	862	$UHUO_{NC}$ = Non-contact, unknown if household/unknown other
4,356	1,596	Voice mail
23	7	Other non-contact (deaf/disabled)
4,379	1,603	UO_{NC} = Non-contact, unknown eligibility
3,061	705	Refusals
752	135	Callbacks
3,813	840	UO_R = Refusal, unknown if eligible
27	35	O = Other (language)
----	5	Child's cell phone
280	45	Out of state
21	1	County out of range
301	51	SO = Screen out
130	31	R = Refusal, known eligible (breakoffs and qualified CBs)
811	193	I = Completed interviews
14,713	5,216	T = Total numbers sampled

continued…

Table 2. Sample Disposition (continued)

68.5%	63.2%	$e1 = (I+R+SO+O+UO_R+UO_{NC})/(I+R+SO+O+UO_R+UO_{NC}+OF+NWC)$ - Est. frame eligibility of non-contacts
75.8%	81.5%	$e2 = (I+R)/(I+R+SO)$ - Est. screening eligibility of unscreened contacts
50.5%	34.9%	$CON = [I + R + (e2*[O + UO_R])]/[I + R + (e2*[O + UO_R + UO_{NC}]) + (e1*e2*UHUO_{NC})]$
21.1%	20.6%	$COOP = I/[I + R + (e2*[O + UO_R])]$
10.6%	**7.2%**	**AAPOR RR3$=I/[I+R+[e2*(UO_R+UO_{NC}+O)]+[e1*e2*UHUO_{NC}]]$ = CON*COOP**