Methods Report for Vanderbilt University's:

Tennessee Horserace Poll 2018

Survey of TN Registered Voters Age 18+

October 16, 2018

Prepared by:
Chintan Turakhia, Jonathan Best, and Jennifer Su
1 Braxton Way
Suite 125
Glen Mills, PA 19342

@ssrs_research | 484.840.4300 | info@ssrs.com

CONTENTS

Note: SSRS is proud to be a Charter Member of the American Association for Public Opinion Research's (AAPOR) *Transparency Initiative*. We support and encourage transparency in the reporting of research results and are committed to AAPOR's Code of Professional Ethics and Practices.

OVERVIEW

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. These polls provide point-in-time data to find out what registered voters in Tennessee think about national, state, and local public policy issues. CSDI at Vanderbilt commissioned SSRS of Glen Mills, PA to conduct the Center's Tennessee Horserace Poll 2018 before the midterms elections.

The Tennessee Horserace Poll 2018 obtained telephone interviews with a representative sample of 800 registered voters, age 18 or older, living in Tennessee. Telephone interviews were conducted by landline (560) and cell phone (240, including 180 without a landline phone). Interviews were done in English from October 8-13, 2018. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 4.9 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 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.

SAMPLING METHODS

The target population for this poll was Tennessee registered voters age 18 or older. SSRS used a registration-based sampling (RBS) approach. RBS sample was 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 eight strata: [1] 18-29 with landline; [2] 30-44 with landline; [3] 45-64 with landline; [4] 65+ with landline; [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.

The sample size for this poll was n=800 interviews. Overall, 70% of respondents were reached via landline and 30% of respondents were reached via cellular telephone. The majority of the voter list is comprised of landline telephone records.

CONTACT PROCEDURES

For respondent selection for landline sample, interviews were conducted with the youngest adult 18+ male/female 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.

DATA COLLECTION, PROCESSING, AND INTEGRATION

Survey Administration

The field period for this study was October 8-13, 2018. 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, 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.

In order 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 in order to locate any errors. Minimal back-coding was done for Question RACE to code open-end responses into prelisted categories where appropriate. No other coding was done for open-end responses.

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 sample was weighted in two stages. The first-stage of the weighting was the application of a base weight to account for different selection probabilities. In the second stage of weighting, sample demographics were matched to population parameters.

Base weight

The sample frame was divided into strata based on age, and younger registered voters were sampled at higher rates in order to get more young respondents. The first step of the weighting was to compute a base weight to

account for the disproportionate sampling across strata. Base weights were computed so that the base weighted distribution of the sample drawn across strata was the same as the distribution of the entire sample frame across the strata. The baseweight for cases in stratum i can be expressed as:

$$BW_i = \frac{N_i}{n_i}$$

Where N_i is the size of stratum i and n_i is the amount of sample drawn from stratum i.

Post-stratification

The second stage of weighting balanced sample demographics to population parameters. The sample was balanced by form to match Tennessee registered voter parameters for sex, age, race, education, region, and party ID. The basic weighting parameters came from a special analysis of the Current Population Survey, November 2018: Voting and Registration Supplement. The region parameter came from the U.S. Census Bureau's 2011-2015 5-Year American Community Survey (ACS) data. The party ID parameter came from an analysis of recent combined SSRS Omnibus data.

The following table lists the dimensions used in the raking.

Table 1: Raking Dimensions

Dimension	Value Label
Sex	Male
	Female
Age	18-29
	30-34
	35-44
	45-54
	55-64
	65+
Race/Ethnicity	White, not Hispanic
	Other
Education	HS graduate or less
	Some college/Associates degree
	College graduate
Region	East
	Nashville
	Central
	Memphis/West
Party ID	Democrat
	Republican
	Independent / Other

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 target population. Table 2 compares weighted and unweighted total sample distributions to population parameters.

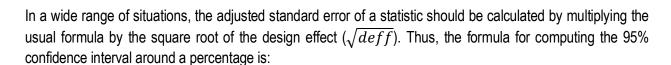### Table 2: Population Parameters and Weighted and Unweighted Total Sample Distributions

Characteristic	Value Label	Parameter	Unweighted	Weighted
Sex	Male	46.3	43.4	45.2
	Female	53.7	56.6	54.8
Age	18-29	14.7	4.0	11.2
	30-34	8.5	2.0	7.8
	35-44	15.8	11.3	17.1
	45-54	16.2	15.6	17.6
	55-64	20.3	28.3	22.3
	65+	24.5	38.9	24.0
Race/Ethnicity	White/not Hispanic	80.0	84.5	79.9
	Other	20.0	15.5	20.1
Education	HS Grad or less	41.6	23.4	37.3
	Some College/Assoc. Degree	24.2	29.6	25.2
	College Graduate	34.2	47.0	37.4
Region	East	36.9	37.0	34.7
	Nashville	23.5	25.1	25.0
	Central	20.4	21.8	21.1
	Memphis/West	19.2	16.1	19.2
Party ID	Democrat	24.2	25.4	24.9
	Republican	30.1	39.9	30.0
	Independent/ Other	45.0	33.9	42.5
	DK/Ref	0.7	0.9	2.6

Effects of Sample Design on Statistical Analysis

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 2.03.

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

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

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:

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

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 ± 4.9 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 4.9 percentage points away from their true values in the population. The margin of error based on Form 1 or 2 respondents is ± 7.3 percentage points. 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 3 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.[1]

- The response rate for RBS landline was 7 percent.
- The response rate for RBS cell was 5 percent.

---

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

Table 3: Sample Disposition

RBS LL	RBS CELL	TOTAL	
90	94	184	Non-residential/Business
12	16	28	Over quota
0	1	1	Cell in landline frame / LL in Cell frame
1	4	5	Duplicate number
103	115	218	OF = Out of Frame
3,602	710	4,312	Not working
333	3	336	Computer/fax/modem
3,935	713	4,648	NWC = Not working/computer
4,098	1,628	5,726	NA/Busy all attempts
4,098	1,628	5,726	$UHUO_{NC}$ = Non-contact, unknown if household/unknown other
2,046	2,632	4,678	Voice mail
819	71	890	Privacy Manager
11	0	11	Other non-contact
2,876	2,703	5,579	UO_{NC} = Non-contact, unknown eligibility
1,246	773	2,019	Refusals
1,466	1,068	2,534	Callbacks
2,712	1,841	4,553	UO_R = Refusal, unknown if eligible
7	33	40	O = Other (language)
0	34	34	Child's cell phone
45	69	114	Other ineligible
45	103	148	SO = Screen out (ineligible)
49	22	71	R = Refusal, known eligible (breakoffs and qualified CBs)
560	240	800	I = Completed interviews
14,385	7,398	21,783	T = Total numbers sampled

RBS LL	RBS CELL	TOTAL	
60.7%	85.6%	69.7%	$e1 = (I+R+SO+O+UO_R+UO_{NC})/(I+R+SO+O+UO_R+UO_{NC}+OF+NWC)$ - Est. frame eligibility of non-contacts
93.1%	71.8%	85.5%	$e2 = (I+R)/(I+R+SO)$ - Est. screening eligibility of unscreened contacts
38.6%	35.3%	37.0%	$CON = [I + R + (e2*[O + UO_R])]/[I + R + (e2*[O + UO_R + UO_{NC}]) + (e1*e2*UHUO_{NC})]$
17.8%	14.9%	16.7%	$COOP = I/[I + R + (e2*[O + UO_R])]$
6.9%	**5.3%**	**6.2%**	**AAPOR RR3**$=I/[I+R+[e2*(UO_R+UO_{NC}+O)]+[e1*e2*UHUO_{NC}]] = $**CON*COOP**

DELIVERABLES

SSRS delivered to Vanderbilt University:

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

ABOUT SSRS

SSRS is a full-service market and survey research firm managed by a core of dedicated professionals with advanced degrees in the social sciences. Service offerings include the Omnibus Survey, Probability Panel and other Online Solutions 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. SSRS is research, refined. Visit www.ssrs.com for more information.