2020
NASHVILLE/ DAVIDSON COUNTY SURVEY

Submitted to: CSDI at Vanderbilt University

Prepared by:

Chintan Turakhia

Jennifer Su

Jonathan Best

MAY 2020

ssrs
research. refined.

AAPOR
Transparency Initiative
AMERICAN ASSOCIATION FOR
PUBLIC OPINION RESEARCH

CONTENTS

SUMMARY

Beginning in 2015, the Center for the Study of Democratic Institutions (CSDI) at Vanderbilt University has conducted public opinion polls of Nashville/Davidson County residents to help inform community stakeholders, government officials, academics, the general public, etc. about important issues facing the Nashville Metro area. The 2020 survey marks Vanderbilt University's fifth poll of the Nashville Metro area.

The 2020 Nashville/Davidson County Poll obtained surveys via web and telephone with a representative sample of 1,036 adults, age 18 or older, living in Nashville-Davidson County, Tennessee. Data collection was conducted in English from April 9, 2020 to May 10, 2020.

Implementing an ABS design, 1,036 respondents were reached via mail. Contacts were asked to complete the survey via web or call a toll-free number to complete the survey by phone. In total, 993 surveys were completed via web and 43 surveys were completed via phone.

Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 4.3 percentage points.

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

SAMPLE DESIGN

Target Population

The target population for this poll was non-institutionalized adults age 18 or older living in Davidson County, Tennessee. SSRS used a mixed mode ABS sampling design. Our sampling approach ensured that we obtained a representative sample of the target population by utilizing a full probability design.

Address-Based Sampling (ABS) Design

The sampling frame under this design was the United States Postal Service (USPS) Computerized Delivery Sequence (CDS) File. The CDS File is a computerized file that contains information on all delivery addresses serviced by the USPS. For this poll, SSRS selected a random sample of addresses in Davidson County from the CDS File. We appended a Hispanic surname flag and low education (HS grad or less) flag for possible targeted outbound call effort.

African Americans (AA) and Hispanics are typically under-represented in ABS samples. To help combat this, Davidson County was divided into five substrata based on African American and Hispanic incidences at the CBG (Census Block Group) level. High AA/Hispanic density strata were oversampled relative to the lower density strata. This stratification plan was expected to help bring AA and Hispanic proportion in the sample closer to population incidences.

Table 1: Sample Stratification

	CBGs		Total Pop		Sample Distribution
High AA (50%+ AA)	82	17.3%	112,832	16.6%	26.7%
Medium AA (30-50% AA)	76	16.1%	124,939	18.4%	29.6%
High Hispanic (20%+ Hisp)	54	11.4%	81,355	12.0%	12.8%
Medium Hisp (10-20% Hisp)	38	8.0%	65,703	9.7%	7.8%
Residual	223	47.1%	293,493	43.3%	23.1%
Total	473	100.0%	678,322	100.0%	100.0%

QUESTIONNAIRE AND LETTER DEVELOPMENT

Questionnaire Development

The questionnaire was developed by the Principal Investigators at the Center for the Study of Democratic Institutions (CSDI) at Vanderbilt University in consultation with the SSRS project team. SSRS reviewed the questionnaire primarily to identify potential problems in the instrument that might increase respondent burden, cause respondents to refuse or terminate the interview, create problems with respondent comprehension, or pose practical challenges for mode-specific administration such as complex skip patterns.

Letter and Postcard Design

The text for the study invitation letter, reminder postcard, and correction follow-up letters were developed by SSRS in consultation with Vanderbilt University investigators. SSRS sent Vanderbilt University the final mailing materials for approval before printing and mailing the materials to contacts.

CONTACT PROCEDURES

Original Procedures

SSRS's original procedures called for a sequential Web-Phone mixed-mode methodology for this study. All sampled households would receive an initial invitation letter and a postcard reminder, with a non-response follow-up mailing. Our planning included completion of a total of up to n=1,000 surveys (n=850 surveys via web mode, n=100 surveys via inbound phone mode and n=50 surveys via outbound phone mode). The outbound phone effort was intended to target the hard-to-reach lower education and Hispanic respondents based on auxiliary data appended to the sample file.

Due to an abundance of web surveys completed within the first few days following the mailing of the invitation letters and postcard reminders, a revised set of contact procedures were enacted as described below.

Revised Procedures

All selected sample records received a one-page, single-sided study invitation letter, which was printed on Vanderbilt University CSDI stationery and was addressed to "Nashville Area Resident." The text of the letter was printed in English only. The text of this letter was developed in collaboration with Vanderbilt University investigators, and it included a short web link for the survey and a personalized PIN for respondents to use to access the web survey. The prenotification mailing also included a $1 pre-incentive bill and an offer of $10 payment via an electronic gift card code to be sent upon completion of the web survey and via check to respondents completing the survey via phone. The letter included a toll-free number for respondents to call to complete the survey via phone if they do not have internet access or preferred to complete the survey via phone.

Two days after mailing of the study invitation letter, reminder postcards were sent to contacts. The purpose of this mailing was simply to remind an eligible adult in target households to reply to the initial mailing. To ensure confidentiality, the postcard did not contain the survey web link or the unique PIN for the web survey.

A non-response follow-up mailing was planned for approximately two weeks after the study invitation letters were mailed. Due to sufficient response from the invitation mailing and postcard reminder to reach the target sample size of 1,000 surveys, the reminder mailing was canceled.

Table 2: Contact Schedule

Date	Mailing	Count
April 9, 2020	Invitation letters mailed	11,765
April 13, 2020	Reminder postcards mailed	11,765
April 23, 2020	Survey packets mailed to select non-responders	Canceled

Correction Mailings

On April 14, 2020, it was discovered that there was an error in the survey program that resulted in respondents not receiving the geographic and age eligibility screening questions. The program was corrected for the remainder of data collection.

Additionally, a re-contact effort was executed to obtain the missing information. Up to two correction mailings were sent to the n=523 respondents affected by the program error, asking the target respondent to answer the two questions they were not originally asked: county of residence and age. Mailings were sent on April 21, 2020 and May 1, 2020. Of the affected respondents, 278 completed the re-contact survey and 245 did not.

PROGRAMMING

Prior to the field period, SSRS programmed the study into its Confirmit platform for web and CATI administration in English. The web program was optimized for administration via smartphone or other mobile handheld devices. Extensive checking of the program was conducted to ensure that skip patterns followed the design of the questionnaire.[1] The web program was checked on multiple devices, including desktop computers and handheld mobile devices, and different web browsers in order to ensure consistent and optimized visualization across devices and web browsers. SSRS generated unique survey passwords that were assigned and provided via mail to potential respondents. The web survey was accessed directly by respondents, using their unique passwords. This also gave them the ability to return to their survey later if they chose to suspend their interview.

DATA COLLECTION, PROCESSING, AND INTEGRATION

Web Data Collection

All contacted households were invited to take the survey online, accessing the survey using a unique PIN to avoid duplication of interviews by the same person.

Telephone Data Collection

Invitation letters included a toll-free number for non-internet or internet-reluctant eligible contacts to participate in the study via telephone. If an interviewer was available at the time of the call, the inbound call was routed to the interviewer for an immediate response. If an interviewer was not available at the time of the call, interested contacts were able to leave a voicemail message to request a callback for a phone interview.

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 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.

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 total, 43 phone interviews were completed from inbound requests. Due to sufficient response via web to reach the target sample size of 1,000 interviews, telephone interviewing was suspended on April 17, 2020 and no outbound dialing was conducted.

---

[1] An error in the program was discovered on April 14 and subsequently corrected.

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 Race other-specify text responses. No other coding was done for open-end responses.

Prior to running cross-tabulations, data from web and telephone modes were combined and thoroughly cleaned with a computer validation program written by one of SSRS's data processing programmers. This program established editing parameters in order to locate any errors, including data that did not follow skip patterns, out of range values, and errors in data field locations.

After quality control procedures were carried out, SSRS provided a clean, processed, fully-labeled and weighted final SPSS dataset to Vanderbilt University.

WEIGHTING AND ANALYSIS

The survey data were weighted to be representative of the adult population of Davidson County, Tennessee. A disproportionately-stratified address based sample was drawn from the U.S. Postal Service's Computerized Delivery Sequence File. The weighting was done in two stages. The first stage was the application of a base weight that corrects for different selection probabilities across sample strata and within households. The second stage of weighting balanced sample demographic distributions to known population benchmarks.

Base weight

The base weight, BW, is the product of a sampling weight (SAMPWT) and a household size adjustment (HHSA). The sampling weight corrects for the disproportionate sample design by adjusting the distribution of completed interviews across the five ABS strata to match the distribution of the ABS frame across strata. The sampling weight for the ABS recruits can be expressed as $\textbf{ABS\_SAMPWT} = P_i/p_i$ where P_i is the proportion of the sample frame from in stratum i and p_i is the proportion of completed interviews in stratum i.

Since we interview one adult in every household regardless of how many adults live in the household, each respondent's probability of being sampled is a function of the number of adults in the household. Specifically, in a household with A adults, a respondent is chosen with probability $1/A$. The household size adjustment is the reciprocal of the selection probability, or $HHSA = A$.[2]

The final base weight is the product of the sampling weight and the household size adjustment.

$$BW = SAMPWT \times \text{HHSA}$$

---

[2] We cap A at 3 to reduce the variance of the weights.

Post-stratification

In the second stage of weighting, sample demographics distributions were adjusted to match known population distributions. The sample was balanced to match Davidson County, Tennessee adult population benchmarks for sex, age, education, race/ethnicity, internet use, and sample stratum. The sex, age, education, and race/ethnicity parameters were derived from the U.S. Census Bureau's 2013-2017 5-year American Community Survey (ACS) data. The Internet use benchmark was derived from ACS 2018. The sample strata distribution came from the Census Planning Database based on ACS 2013-2017 5-Year estimates.

The following table lists the dimensions used in the raking.

Table 3: Raking Dimensions

Dimension	Value Label	
Gender by Age	Male, 18-24	Female, 18-24
	Male, 25-34	Female, 25-34
	Male, 35-44	Female, 35-44
	Male, 45-54	Female, 45-54
	Male, 55-64	Female, 55-64
	Male, 65+	Female, 65+
	missing on age	
Sex by Education	Male, High School Grad or less	
	Male, Some College / Associates Degree	
	Male, College Grad or higher	
	Female, High School Grad or less	
	Female, Some College / Associates Degree	
	Female, College Grad or higher	
Race/Ethnicity	White, not Hispanic	
	Black or African-American, not Hispanic	
	Hispanic	
	Other, not Hispanic	
Internet Use	Yes	
	No	
Sample Strata	High AA	
	Medium AA	
	High Hispanic	
	Medium Hispanic	
	Residual	

Raking was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure. Missing values on the weighting variables were

imputed using a Hot Deck procedure except for age. Due to a programming error, age had too many missing cases to effectively impute. Therefore, cases missing an age were kept in their own weighting class and the gender by age benchmarks were re-percentaged accordingly.

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 4 compares weighted and unweighted distributions to population parameters.

Table 4: Population Parameters and Weighted and Unweighted Sample Distributions

Characteristic	Value Label	Parameter	Unweighted	Weighted
Gender by Age	Male, 18-24	4.0%	1.7%	4.0%
	Male, 25-34	9.3%	7.5%	9.7%
	Male, 35-44	6.9%	5.8%	6.9%
	Male, 45-54	5.9%	3.1%	6.0%
	Male, 55-64	5.3%	5.0%	4.0%
	Male, 65+	4.7%	6.8%	4.4%
	Female, 18-24	4.4%	3.6%	4.6%
	Female, 25-34	10.1%	11.8%	10.6%
	Female, 35-44	6.8%	8.2%	7.0%
	Female, 45-54	6.3%	5.6%	6.4%
	Female, 55-64	6.2%	8.2%	6.5%
	Female, 65+	6.5%	9.1%	6.8%
	missing	23.6%	23.6%	23.1%
Gender by Education	Male, HS Grad or less	17.9%	5.3%	15.2%
	Male, Some College	12.1%	9.6%	12.4%
	Male, College+	17.2%	23.3%	18.0%
	Female, HS Grad or less	17.2%	12.7%	17.2%
	Female, Some College	14.7%	16.4%	15.4%
	Female, College+	20.8%	32.7%	21.8%
Race/Ethnicity	White, non-Hispanic	60.9%	62.9%	60.7%
	Black, non-Hispanic	25.6%	24.0%	25.7%
	Hispanic	8.0%	7.0%	8.0%
	Other, non-Hispanic	5.5%	6.0%	5.7%
Internet Use	Yes	93.8%	98.4%	95.6%
	No	6.2%	1.6%	4.4%
Sample Strata	High density AA	15.6%	23.4%	16.0%
	Medium density AA	17.9%	28.5%	18.7%
	High density Hispanic	11.0%	12.9%	11.4%
	Medium density Hispanic	9.7%	7.3%	10.1%
	Residual	45.7%	27.9%	43.8%

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 design effect for the total sample is 2.02.

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

$$deff = \frac{n \sum w^2}{(\sum w)^2} \, [3]$$

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:

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

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 total sample is ±4.3 percentage points. This means that in 95 out of every 100 samples drawn using the same methodology, estimated proportions based on the sample will be no more than 4.3 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.

---

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

RESPONSE RATES

Table 5 reports the disposition of all sampled ABS records that were contacted. 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. The response rate for the ABS sample was 8.8%.

Table 5: ABS Sample Disposition

Disposition	N
1. Complete (I)	**1,036**
2. Eligible, non-interview (R)	**66**
Refusal	3
Break-off	63
3. Unknown eligibility, non-interview (UH)	**10,661**
Nothing ever returned	10,089
Refused, unknown if eligible	4
Over-quota	97
Undeliverable	471
4. Not eligible, returned (IN)	**2**
Currently lives in Davidson County	0
Do not live in Davidson County	2
Total records contacted	**11,765**
e=(I+R)/(I+R+IN)	99.8%
RR3=I/[I+R+(e*UH)]	**8.8%**

DELIVERABLES

SSRS delivered to Vanderbilt University:

- Final questionnaire instrument;
- Weighted dataset in SPSS;
- Weighted banners in PDF;
- Topline; and
- A detailed methodology 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, SSRS Opinion 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.