2023
NASHVILLE/ DAVIDSON COUNTY
SURVEY

Submitted to: CSDI at Vanderbilt University

Prepared by:

Chintan Turakhia

Jennifer Su

Hayley Boote

APRIL 2023

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 2023 survey marks Vanderbilt University's 8th poll of Metro Nashville.

The 2023 Nashville/Davidson County Poll obtained surveys via web and telephone with a representative sample of 1,016 adults, ages 18 or older, living in Nashville-Davidson County, Tennessee. Data collection was conducted in English from March 13 to April 6, 2023.

Implementing an Address-Based Sample (ABS) design, 1,016 respondents were first contacted 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, 958 surveys were completed via web and 58 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 ± 4.1 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 ages 18 or older living in Davidson County, Tennessee. SSRS used an Address-Based Sampling (ABS) design for this mixed mode study. Our sampling approach ensured that we obtained a representative sample of the target population via 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. Sample was provided by SSRS sister company Marketing Systems Group (MSG). 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 mailing effort.[1]

Hard-to-reach demographic groups such as blacks/African Americans, Hispanics, and those with lower educational attainment are less likely to participate in surveys. For this poll, SSRS combined the successful stratification used for the 2022 Nashville survey with predictive models created by the SSRS Methods,

---

[1] No targeted mailing effort was conducted.

Analytics, and Data Science (MADS) group to account for the under-representation of these groups. As was done for the 2022 Nashville Survey, we implemented a stratified design that sampled Census Block Groups (CBGs) with higher concentrations of African American (AA) or Hispanic residents at a slightly higher rate to ensure adequate African American and Hispanic representation in the sample. In addition to the census-based stratification, SSRS utilized predictive models specifically designed to identify addresses with a higher probability of having an adult household member with less than a high school education (LTHS). These models were built by appending commercially available data (at the address level) to the data collected from the 2022 Nashville Study. This ensured the applicability of the models to the specific study. These models were then applied to the new ABS sample such that addresses likely to include an adult household member with less than a high school education could be sampled at a slightly higher rate to increase representativeness.

Table 1A: ABS Frame Counts

Universe	n	%
High AA 2010 Block Groups	50,931	15.1%
Medium AA 2010 Block Groups	58,723	17.4%
High Hispanic 2010 Block Groups	32,466	9.6%
Medium Hisp 2010 Block Groups	31,516	9.3%
Residual 2010 Block Groups	163,950	48.6%
Total	337,586	100.0%

Table 1B: Sample Stratification and Distribution

Sample Distribution	n	%
High AA (50%+ AA), LTHS	292	2.9%
High AA (50%+ AA), Not LTHS	2,522	25.2%
Medium AA (30-50% AA), LTHS	365	3.7%
Medium AA (30-50% AA), Not LTHS	2,822	28.2%
High Hispanic (20%+ Hisp.), LTHS	109	1.1%
High Hispanic (20%+ Hisp.), Not LTHS	834	8.3%
Medium Hisp. (10-20% Hisp.), LTHS	62	0.6%
Medium Hisp. (10-20% Hisp.), Not LTHS	825	8.3%
Residual, LTHS	45	0.5%
Residual, Not LTHS	2,124	21.2%
Total	10,000	100.0%

For the selection of ABS records, the following addresses filters were imposed:

- Inclusions:
  - o <u>Address Types</u>: Residential (single family dwelling or multi-family dwelling), Residential with some business.
  - o <u>Primary Address Filters</u>: City style, Rural route (a non-city style address), Highway contract route (routes located on the highway that are contracted out for delivery), Only Way to Get Mail PO Boxes (OWGM).
  - o <u>Secondary Address Filters</u>: Throwback (address is an actual street address, but delivery is made to a customer's PO Box), Drop points (central mail receptacle for multiple units, and limited, if any, information in the ABS frame at the unit level, such as apartment or unit numbers).
- Exclusions:
  - o <u>Address Types</u>: Business, Business with some residential.
  - o <u>Primary Address Filters</u>: Traditional PO Boxes (box holders who have a deliverable street address but choose to have their mail delivered to a PO Box and therefore could be sampled more than once).
  - o <u>Secondary Address Filters</u>: Seasonal (addresses where mail is delivered only during a specific season such as a summer residence), Educational (non-dormitory housing units near the campus of an academic institution), Vacant (delivery points that have been unoccupied for 90 days or longer), No Stat (addresses not currently receiving mail delivery [for example, planned development] and are not in the CDS).

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 reminder letter 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 sampled households.

CONTACT PROCEDURES

Two-Release Approach

To ensure all survey programs and data collection systems were operating as expected, SSRS implemented a two-release approach for the contact protocol. Release #1 began with a sample of 2,500 households. After some surveys had been completed and the system thoroughly tested in the live environment, an additional 7,500 pieces of sample were released in the second wave. Release #2 followed the same general contact protocol established for Release #1.

Description of Mailings

All sampled households were first mailed an initial invitation letter. This invitation letter was one-page, single-sided, and printed in English. This letter was printed on Vanderbilt University's stationery (produced by SSRS). The text of the letter was developed in collaboration with Vanderbilt University investigators and include a short web link for the survey and a unique password to access the secure web survey. The letter also included a $1 bill as a pre-incentive and an offer of $10 payment after completion of the survey. Inclusion of a pre-incentive generally improves the overall study participation rate by acting as a token of appreciation and encouraging participation in the survey. To maximize response by offering invitees more than one way of completing the survey, this letter included a toll-free number that potential respondents could call to complete the survey via phone if they did not have internet access or preferred to complete the survey via phone.

Approximately 2-3 business days after mailing the invitation letter, a reminder postcard was sent to all sampled households. The purpose of this mailing was simply to remind an eligible adult in the household to look for the initial mailing and complete the survey. The postcard was an open-face (not sealed) card and did not contain the survey web link or the respondent's unique password to preserve respondent confidentiality.

Approximately two weeks after the study invitation letters were mailed, a reminder letter was mailed to all non-responders. This letter stressed the urgency of this final contact and included a toll-free number that potential respondents could call to complete the survey via phone.

All mailings were sent via USPS first-class mail.

Table 2: Contact Schedule

Date	Mailing	Count
March 13, 2023	Release #1 invitation letters mailed	2,500
March 15, 2023	Release #1 postcards mailed	2,500
March 16, 2023	Release #2 invitation letters mailed	7,500
March 20, 2023	Release #2 postcards mailed	7,500
March 30, 2023	Release #1 reminder letters mailed	2,387
March 30, 2023	Release #2 reminder letters mailed	7,109

Post-compensation

For those completing the web survey, a $10 via an electronic gift code was sent to each respondent immediately after the survey, enabling them to use the gift card as soon as they received it. For those completing via phone, a check was mailed. Mailed post-incentives were sent on a weekly basis during data collection.

DATA COLLECTION, PROCESSING, AND INTEGRATION

Web Data Collection

All contacted households were invited to take the survey online, accessing the survey using a unique password 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, 34 phone interviews were completed from inbound requests and 24 phone surveys were completed from outbound dialing.

Targeted Outbound Dialing

Several days into the field period, SSRS began conducting targeted outbound dialing for harder-to-reach populations, specifically lower-educated adults (high school graduate or equivalent, or less education) and Hispanics. Targeted outbound dialing was limited to these groups because they are typically under-represented in general population public opinion surveys. Outbound dialing for both batches of sample began on March 29.

PROGRAMMING, DATA PROCESSING, AND INTEGRATION

Programming

Prior to the field period, SSRS programmed the study into its Forsta Plus (formerly known as Confirmit) platform for web and CATI administration in English. The web program was optimized for administration via smartphone or other handheld mobile devices. Extensive checking of the program was conducted to ensure that skip patterns followed the design of the questionnaire. The web program was checked on multiple devices, including desktop computers and handheld mobile devices, and different web browsers to ensure consistent and optimized visualization across devices and web browsers. SSRS generated unique survey passwords that were assigned at the sample level 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.

Quality Control Checks

Quality checks were incorporated into the survey. Respondents who failed the following quality checks were not included in the final dataset:

1. Web respondents who answered two or more trap questions incorrectly (n=4).
2. Respondents who answered less than 90% of the questions asked of them if web or answered less than 70% of the questions asked of them if phone (n=0).
3. Respondents who completed the survey in less than 30% of the median length[2] (n=7).

A total of N=11 completed surveys were removed after applying these cleaning standards.

---

[2] A speeder was a complete achieved in 2.55 minutes or less.

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.

SSRS also reviewed item non-response and completion time to evaluate the data quality for each survey. Three surveys with about 20%+ non-response and one survey completing in less than three minutes were rejected and removed from the data provided to Vanderbilt University.

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

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 sample was weighted in stages. The first stage of the weighting was the application of a base weight to account for different selection probabilities and response rates across sample strata. In the second stage sample demographics were post-stratified to match population parameters.

Base weight

The base weight begins with a stratification weight for all sampled addresses that adjusts for disproportionate sampling across the ten design strata, defined as five geographic strata (High AA, Medium AA, High Hispanic, Medium Hispanic, Residual) crossed by two modeled low education strata (low education, not low education). It is expressed as $SAMPWT_{i,h} = N_h/n_h$ where N_h is the size of the sample frame in stratum h and n_h is the number of addresses mailed in stratum h.

The stratification weights were then adjusted for household eligibility and nonresponse. Eligibility and nonresponse adjustments were conducted together. To implement the eligibility/nonresponse adjustment, the sample was divided into nonresponse adjustment cells defined by the five geographic sampling strata. Within each cell c, the nonresponse adjustment factor was calculated as:

$$NRADJ_c = \frac{\sum_{c,sample-ineligible} SAMPWT}{\sum_{c,complete} SAMPWT}$$

That is, the sum of the design weights among all sampled cases within the cell (excluding ineligible addresses), divided by the sum of the sampling weights among all cases with a completed survey within the cell.

After applying the nonresponse adjustment, the weights were then adjusted for the sampling of one adult in each sampled household. The probability of selection adjustment (PSA_i) for each respondent i is the number of adults in the respondent household.[3]

The final adjusted base weight is defined for completes only and incorporates the nonresponse and probability of selection adjustment factor.

$$BW_{i,h,c} = \begin{cases} SAMPWT_{i,h} \times NRADJ_c \times PSA_i, & if\ complete \\ 0, & otherwise \end{cases}$$

Post-stratification

With the final base weight applied, the data were weighted to balance the demographic profile of the sample to the target population parameters. The sample was weighted to match population estimates. Benchmarks were derived from 5-year American Community Survey[4] (ACS) data. The weighting parameters were gender by age, gender by education, race/ethnicity, internet access status, and geographic sampling strata.

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+
Gender by Education	Male, High School Grad or less	
	Male, Some College / Associate Degree	
	Male, College Grad or higher	
	Female, High School Grad or less	
	Female, Some College / Associate Degree	
	Female, College Grad or higher	
Race/Ethnicity	White, not Hispanic	
	Black or African American, not Hispanic	
	Hispanic	

---

[3] For this adjustment, the number of adults will be capped at three to contain the variance of the weights.
[4] https://data.census.gov/table

	Other, not Hispanic
Internet Access	Access
	No access
	High African American
	Medium African American
Geographic Strata	High Hispanic
	Medium Hispanic
	Residual

Weighting was accomplished using the R package ANESRAKE.[5] 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).

The following table compares weighted and unweighted distributions to population parameters.

[5] https://cran.r-project.org/web/packages/anesrake/anesrake.pdf

Table 4: Population Parameters and Weighted and Unweighted Sample Distributions

Category	Values	Parameter	Unweighted	Weighted
Gender by Age	Male 18-24	4.9%	2.2%	3.8%
	Male 25-34	12.3%	10.5%	11.7%
	Male 35-44	9.2%	5.9%	9.5%
	Male 45-54	7.1%	5.3%	7.3%
	Male 55-64	6.7%	6.2%	6.8%
	Male 65+	7.2%	10.6%	7.4%
	Female 18-24	5.1%	5.1%	5.3%
	Female 25-34	13.7%	15.8%	14.1%
	Female 35-44	9.4%	11.7%	9.8%
	Female 45-54	7.3%	7.0%	7.6%
	Female 55-64	7.6%	9.4%	7.2%
	Female 65+	9.5%	10.2%	9.6%
Gender by Education	Male High School or Less	16.1%	6.3%	13.7%
	Male Some College	10.2%	9.4%	10.6%
	Male College+	21.2%	25.0%	22.1%
	Female High School or Less	15.5%	10.1%	15.0%
	Female Some College	13.0%	15.0%	13.5%
	Female College+	24.0%	34.2%	25.0%
Race/Ethnicity	White, non-Hispanic	59.8%	61.9%	60.0%
	Black, non-Hispanic	24.2%	23.7%	23.8%
	Hispanic	8.3%	8.8%	8.7%
	Other, non-Hispanic	7.6%	5.6%	7.6%
Internet Access	Access	95.8%	97.8%	96.5%
	No access	4.2%	2.2%	3.5%
Geographic Strata	High AA	15.5%	24.5%	16.1%
	Medium AA	18.2%	29.3%	18.7%
	High Hisp	10.2%	7.8%	10.1%
	Medium Hisp	9.5%	10.0%	9.5%
	Residual	46.7%	28.3%	45.6%

Weights were trimmed at the 2nd and 98th percentiles to prevent individual interviews from having too much influence on survey-derived estimates. 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.

Margin of Sampling Error

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 1.77.

SSRS calculates the composite design effect for a sample of size n, with each case having a weight, w, as:[6]

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

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 sample is ±4.1 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 4.1 percentage points away from their true values in the population. Margins of error for subgroups will be larger.

It is important to remember that the sampling fluctuations captured in the margin of error 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.

---

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

RESPONSE RATES

The following table reports the disposition of all sampled ABS records that were contacted. The response rate estimates the fraction of all eligible sample that was ultimately surveyed. Response rates are computed according to American Association for Public Opinion Research (AAPOR) standards. The response rate for the ABS sample was 15.5%.

Table 5: ABS Sample Disposition

Disposition	N
Complete (I)	**1,016**
Web	958
Phone	58
Eligible, non-interview (R)	**5**
Refusal & Break-off	5
Unknown eligibility, non-interview (UH)	**8,442**
Nothing ever returned	8,071
Refused, unknown if eligible	364
Over-quota	7
Not eligible, returned (IN)	**537**
Does not live in Davidson County, TN	23
Not age 18 or older	6
Business/Commercial	11
Undeliverable	497
Total records contacted	**10,000**
e=(I+R)/(I+R+IN)	65.5%
RR3=I/[I+R+(e*UH)]	**15.5%**

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 Panel, 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. SSRS is research, refined. Visit www.ssrs.com for more information.