

Methodology

2016 Indiana Voter Survey

Prepared by Princeton Survey Research Associates International
for Ball State University

October 2016

The 2016 Indiana Voter Survey obtained telephone interviews with a representative sample of 803 registered voters in Indiana. Telephone interviews were conducted by landline (642) and cell phone (161). The survey was conducted by Princeton Survey Research Associates International (PSRAI). Interviews were done in English by Princeton Data Source from October 10 to October 16, 2016. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error based on the entire sample is ± 3.9 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 Indiana State voter file. The database has approximately 4.3 million records in total including approximately 1.1 million with a phone number. The sample frame that included a phone number was split into four strata: [1] 18-44 with landline; [2] 18-44 with only cell phone; [3] 45+ with landline; and [4] 45+ 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 October 10 to October 16, 2016. 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 cellular respondents were in a safe place before administering the survey.

Once the target respondent was on the phone, interviewers confirmed that they lived in Indiana 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 Indiana registered voter parameters.

The weighting balanced sample demographics to population parameters. The final sample was balanced to match Indiana registered voter parameters for sex, age, race, education and region¹. The sex, age and region parameters were provided by Aristotle's Demographic Profile of Registered Voters in IN. The race and education parameters came from an analysis of the November 2015 survey data that PSRAI conducted for Ball State.

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 Indiana counties were divided into three regions. The regions were Northern, Central and Southern.

Table 1. Sample Demographics

	<u>Parameter</u>	<u>Unweighted</u>	<u>Weighted</u>
<u>Sex</u>			
Male	47.4	44.5	46.9
Female	52.6	55.5	53.1
<u>Age</u>			
18-24	7.8	5.2	7.9
25-34	16.9	9.8	15.8
35-44	16.9	16.3	17.2
45-54	18.0	12.4	17.9
55-64	18.3	16.4	18.6
65+	22.1	39.9	22.5
<u>Race/Ethnicity</u>			
White, not Hispanic	85.6	89.4	86.1
Other	14.4	10.6	13.9
<u>Region</u>			
Northern	32.9	32.8	33.0
Central	43.1	43.2	42.9
Southern	24	24.0	24.1
<u>Education</u>			
HS grad or less	39.8	32.0	38.8
Some college /Assoc.	33.8	29.4	34.3
4-yr degree+	26.4	38.5	26.9

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 systematic non-response. The total sample design effect for this survey is 1.26.

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

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

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 ± 3.9 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 3.9 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 likely voters is ± 4.8 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 telephone numbers ever dialed from the original telephone number samples. The response rate estimates the fraction of all eligible sample that was ultimately interviewed. Sample disposition reporting and response rates calculations are in accordance with American Association for Public Opinion Research standards.² Thus the response rates were 14 percent for the landline sample and 19 percent for the cell sample.

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

Self-Identified Party Identification (Likely Voters)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Republican	123	22.6	22.6	22.6
	Leans Republican	81	14.9	14.9	37.5
	Independent	122	22.4	22.4	59.9
	Leans Democrat	86	15.8	15.8	75.7
	Democrat	115	21.1	21.1	96.9
	No party/Not interested in politics (VOL.)	10	1.8	1.8	98.7
	Other party (SPECIFY) (VOL.)	2	.4	.4	99.1
	Don't know/Refused (VOL.)	5	.9	.9	100.0
	Total	544	100.0	100.0	

In Indiana, voters do not register by party. Many self-identify as Independent but may vote as partisans.

Table 2. Sample Disposition

<u>Landline</u>	<u>Cell</u>	
118	21	Non-residential/Business
1	0	Cell in landline frame
119	21	OF = Out of Frame
1,712	334	Not working
144	2	Computer/fax/modem
1,856	336	NWC = Not working/computer
738	66	NA/Busy all attempts
0	389	VM not set up/caller out of range
738	455	UHUO _{NC} = Non-contact, unknown if household/unknown other
5,157	1,570	Voice mail
49	9	Other non-contact (deaf/disabled/deceased)
5,206	1,579	UO _{NC} = Non-contact, unknown eligibility
1,437	254	Refusals
904	163	Callbacks
2,341	417	UO _R = Refusal, unknown if eligible
12	13	O = Other (language)
732	433	Listed person does not exist
0	2	Child's cell number
74	33	Out of state
806	468	SO = Screen out
57	16	R = Refusal, known eligible (breakoffs and qualified CBs)
642	161	I = Completed interviews
11,777	3,466	T = Total numbers sampled
82.1%	88.1%	$e1 = (I+R+SO+O+UO_R+UO_{NC}) / (I+R+SO+O+UO_R+UO_{NC}+OF+NWC)$ - Est. frame eligibility of non-contacts
46.4%	27.4%	$e2 = (I+R) / (I+R+SO)$ - Est. screening eligibility of unscreened contacts
39.9%	35.2%	$CON = [I + R + (e2*[O + UO_R])] / [I + R + (e2*[O + UO_R + UO_{NC}]) + (e1*e2*UHUO_{NC})]$
35.8%	54.6%	$COOP = I / [I + R + (e2*[O + UO_R])]$
14.3%	19.2%	AAPOR RR3 = $I / [I + R + [e2*(UO_R + UO_{NC} + O)] + [e1*e2*UHUO_{NC}]] = CON*COOP$

