Methodology

2010 Hoosier Poll

Prepared by Princeton Survey Research Associates International
for Ball State University

November 2010

SUMMARY

The 2010 Hoosier Poll, sponsored by Ball State University, obtained telephone interviews with a representative sample of 601 adults living in Indiana. Telephone interviews were conducted by landline (420) and cell phone (181, including 73 with adults with no landline phone). The survey was conducted by Princeton Survey Research Associates International (PSRAI). Interviews were done in English by Princeton Data Source from November 9-18, 2010. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ±4.7 percentage points.

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

DESIGN AND DATA COLLECTION PROCEDURES

Sample Design

A combination of landline and cellular random digit dial (RDD) samples was used to represent all adults in Indiana who have access to either a landline or cellular telephone. Both samples were provided by Survey Sampling International, LLC (SSI) according to PSRAI specifications.

Numbers for the landline sample were drawn with equal probabilities from active blocks (area code + exchange + two-digit block number) that contained three or more residential directory listings. The cellular sample was not list-assisted, but was drawn through a systematic sampling from dedicated wireless 100-blocks and shared service 100-blocks with no directory-listed landline numbers.

Contact Procedures

Interviews were conducted from November 9 to November 18, 2010. As many as seven attempts were made to contact every landline telephone number and as many as five attempts were made to each cell 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 youngest adult male or female currently at home based on a random rotation. If no male/female was available, interviewers asked to speak with the youngest adult of the other gender. This systematic respondent selection technique has been shown to produce samples that closely mirror the population in terms of age and gender when combined with cell interviewing. The interviewers also confirmed the respondent was a current resident of Indiana.

For the cellular sample, interviews were conducted with the person who answered the phone. Interviewers verified that the person was an adult currently residing in Indiana and in a safe place before administering the survey.

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 sample was weighted to match adult general population parameters for Indiana. A two-stage weighting procedure was used to weight this dual-frame sample.

The first stage of weighting corrected for different probabilities of selection associated with the number of adults in each household and each respondent's telephone usage patterns.[1] This weighting also adjusts for the overlapping landline and cell sample frames and the relative sizes of each frame and each sample.

---

[1] i.e., whether respondents have only a landline telephone(s), only a cell phone, or both kinds of telephone.

This first-stage weight for the ith case can be expressed as:

$$WT_i = \frac{1}{\left(\frac{S_{LL}}{S_{CP}} \times \frac{1}{AD_i}\right)} \; if \; respondent \; has \; no \; cell \; phone$$

$$WT_i = \frac{1}{\left(\frac{S_{LL}}{S_{CP}} \times \frac{1}{AD_i}\right) + R} \; if \; respondent \; has \; both \; kinds \; of \; phones$$

$$WT_i = \frac{1}{R} \; if \; respondent \; has \; no \; land \; line \; phone$$

Where S_{LL} = size of the landline sample

S_{CP} = size of the cell phone sample

AD_i = Number of adults in the household

R = Estimated ratio of the land line sample frame to the cell phone sample frame

The equations can be simplified by plugging in the values for S_{LL} = 420 and S_{CP} = 181. Additionally, we will estimate of the ratio of the size of landline sample frame to the cell phone sample frame R = 1.08.

$$WT_i = \frac{1}{\left(\frac{420}{181} \times \frac{1}{AD_i}\right)} \; if \; respondent \; has \; no \; cell \; phone$$

$$WT_i = \frac{1}{\left(\frac{420}{181} \times \frac{1}{AD_i}\right) + 1.08} \; if \; respondent \; has \; both \; kinds \; of \; phones$$

$$WT_i = \frac{1}{1.08} \; if \; respondent \; has \; no \; land \; line \; phone$$

The second stage of weighting balanced sample demographics to population parameters. The sample is balanced to match parameters for sex, age, education, race, Hispanic origin, region[2], population density, and telephone usage. The basic weighting parameters came from a special analysis of the Census Bureau's 2009 Annual Social and Economic Supplement (ASEC) that included all households in Indiana State. The population density and region parameters were derived from Census 2000 data. The telephone usage parameter came an analysis of recently conducted PSRAI surveys, looking at the phone use patterns for Indiana respondents. Those

---

[2] Region was broken into three groups using the FIPS county codes. The groups were "Northern", "Central", and "Southern" Indiana.

original surveys were weighted to account for telephone usage patterns derived from of the July-December 2009 National Health Interview Survey.[3]

Weighting was accomplished using Sample Balancing, a special iterative sample weighting program that simultaneously balances the distributions of all variables using a statistical technique called the *Deming Algorithm*. 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.

---

[3] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December, 2009. National Center for Health Statistics. May 2010.

Table 1: Sample Demographics

	Parameter	Unweighted	Weighted
Gender			
Male	48.5	43.1	48.8
Female	51.5	56.9	51.2
Age			
18-24	12.1	6.3	11.6
25-34	17.2	9.0	15.3
35-44	18.4	15.1	18.2
45-54	17.6	21.0	18.1
55-64	17.0	20.3	17.9
65+	17.7	28.3	18.9
Education			
Less than HS Graduate	12.8	7.5	11.7
HS Graduate	43.8	35.7	43.4
Some College	22.9	26.0	23.1
College Graduate	20.5	30.7	21.9
Race/Ethnicity			
White/not Hispanic	87.5	89.2	88.2
Black/not Hispanic	7.2	7.6	7.3
Hispanic/Other	5.3	3.2	4.4
Region			
Northeast	33.9	33.6	33.9
Central	42.5	44.1	42.1
Southern	23.6	22.3	23.9
County Pop. Density			
1 - Lowest	19.5	21.5	20.2
2	32.2	31.8	32.6
3	26.1	28.6	26.1
4	8.0	6.5	7.8
5 - Highest	14.2	11.6	13.4
Household Phone Use			
LLO	13.0	11.1	12.3
Dual	65.0	76.7	66.5
CPO	22.0	12.1	21.2

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 1.36.

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

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

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:

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

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 ±4.7 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 4.7 percentage points away from their true values in the population. 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 report 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. At PSRAI it is calculated by taking the product of three component rates:[4]

- o Contact rate – the proportion of working numbers where a request for interview was made[5]
- o Cooperation rate – the proportion of contacted numbers where a consent for interview was at least initially obtained, versus those refused
- o Completion rate – the proportion of initially cooperating and eligible interviews that were completed

Thus the response rate for the land line samples was 24 percent. The response rate for the cellular sample was 21 percent.

---

[4] PSRAI's disposition codes and reporting are consistent with the American Association for Public Opinion Research standards.
[5] PSRAI assumes that 75 percent of cases that result in a constant disposition of "No answer" or "Busy" are actually not working numbers.

Table 2: Sample Disposition

Landline	Cell	
6,255	2,699	Total Numbers Dialed
264	48	Non-residential
221	1	Computer/Fax
0	----	Cell phone
3,635	1,113	Other not working
192	25	Additional projected not working
1,943	1,512	Working numbers
31.1%	56.0%	Working Rate
64	8	No Answer / Busy
318	336	Voice Mail
7	1	Other Non-Contact
1,554	1,167	Contacted numbers
80.0%	77.2%	Contact Rate
75	147	Callback
1,001	692	Refusal
478	328	Cooperating numbers
30.8%	28.1%	Cooperation Rate
17	22	Language Barrier
32	119	Not an Indiana Resident/Child's cell phone
429	187	Eligible numbers
89.7%	57.0%	Eligibility Rate
9	6	Break-off
420	181	Completes
97.9%	96.8%	Completion Rate
24.1%	21.0%	Response Rate