Methodology

Massachusetts U.S. Senate Poll

Prepared by Princeton Survey Research Associates International

September 2012

Results for the Massachusetts U.S. Senate Poll are based on telephone interviews with a random sample of 524 Massachusetts registered voters. Telephone interviews were conducted by landline (325) and cell phone (199, including 58 without a landline phone). The survey was conducted by Princeton Survey Research Associates International (PSRAI). Interviews were done in English by Princeton Data Source from September 13-17, 2012. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ±5.3 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 the Massachusetts 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 September 13-17, 2012. 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 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.

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.

All cooperating respondents from both samples were asked about their voter registration status. Registered voters continued with the full interview. People who said they were not registered to vote, or were uncertain about their registration status were only asked demographic questions for weighting purposes.

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 full sample (completed interviews plus non-registered voter screenouts) was weighted to match Massachusetts adult population parameters. A two-stage weighting procedure was used to weight this dual-frame sample. After the weighting, the non-registered voters were dropped from analysis.

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, only a cell phone, or both kinds of telephone.

This first-stage weight for the i[th] 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} = 399 and S_{CP} = 288. Additionally, we will estimate of the ratio of the size of landline sample frame to the cell phone sample frame R = 0.64.

The second stage of weighting balanced sample demographics to population parameters. The sample was balanced to match Massachusetts population parameters for sex, age, education, race, Hispanic origin, region[2], population density, number of adults in household, and telephone usage. The basic weighting parameters came from a special analysis of the Census Bureau's 2011 Annual Social and Economic Supplement (ASEC) that included all households in Massachusetts. The population density variable was derived from 2010 Census data. The telephone usage parameter came from an analysis recent PSRAI Omnibus data.

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 target population. Table 1 compares weighted and unweighted total sample distributions to population parameters.

[2] Massachusetts' counties were divided in six regions based on zip codes. The regions were West, Central, Southeast, Outer suburbs, Inner suburbs, and Boston proper.

Table 1: Sample Demographics

	Parameter	Unweighted	Weighted
Gender			
Male	48.5	47.2	49.3
Female	51.5	52.8	50.7
Age			
18-24	12.7	10.9	13.1
25-34	15.3	11.2	15.0
35-44	18.6	10.9	17.5
45-54	21.0	16.9	21.4
55-64	14.1	21.4	14.6
65+	18.3	28.6	18.3
Education			
Less than HS Graduate	10.0	4.3	8.4
HS Graduate	28.3	24.9	27.8
Some College	23.7	23.7	24.4
College Graduate	38.0	47.1	39.4
Race/Ethnicity			
White/not Hispanic	80.7	76.5	80.3
Black/not Hispanic	5.9	6.0	5.8
Hispanic	7.1	5.7	7.2
Other/not Hispanic	6.3	11.9	6.7
Household Phone Use			
LLO	6.5	7.7	6.3
Dual	66.6	77.0	67.7
CPO	26.9	15.3	25.9
Region			
West	12.6	13.0	12.5
Central	12.8	14.4	13.1
Southeast	13.0	15.9	13.0
Outer suburbs	33.7	34.8	34.5
Inner suburbs	17.0	15.0	16.7
Boston proper	10.9	7.0	10.3
Population Density			
Lowest -1,2	5.9	6.7	5.9
3	30.1	32.9	30.4
4	53.0	53.4	53.4
Highest - 5	10.9	7.0	10.3
# of adults in HH			
One	16.2	31.1	17.0
Two	50.9	46.1	50.5
Three +	32.9	22.7	32.5

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 1.56.

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 of registered voters is ±5.3 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 5.3 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 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. At PSRAI it is calculated by taking the product of three component rates:[3]

- Contact rate – the proportion of working numbers where a request for interview was made[4]
- Cooperation rate – the proportion of contacted numbers where a consent for interview was at least initially obtained, versus those refused
- Completion rate – the proportion of initially cooperating and eligible interviews that were completed

Thus the response rate for the land line samples was 11 percent. The response rate for the cellular samples was 15 percent.

[3] PSRAI's disposition codes and reporting are consistent with the American Association for Public Opinion Research standards.

[4] 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	
9927	5999	**T** Total Numbers Dialed
464	78	**OF** Non-residential
396	3	**OF** Computer/Fax
2	0	**OF** Cell phone
4161	2067	**OF** Other not working
445	183	**UH** Additional projected not working
4459	3668	Working numbers
44.9%	61.1%	Working Rate
148	61	**UH** No Answer / Busy
2087	1380	**UO_{NC}** Voice Mail
8	2	**UO_{NC}** Other Non-Contact
2216	2225	Contacted numbers
49.7%	60.7%	Contact Rate
157	325	**UO_{R}** Callback
1561	1312	**UO_{R}** Refusal
498	588	Cooperating numbers
22.5%	26.4%	Cooperation Rate
81	88	**IN1** Language Barrier
85	290	**IN2** Child's cell phone / Non voter or MA resident
332	210	Eligible numbers
66.7%	35.7%	Eligibility Rate
7	11	**R** Break-off
325	199	**I** Completes
97.9%	94.8%	Completion Rate
10.9%	15.2%	Response Rate