

Methodology

Massachusetts U.S. Senate Registered Voter Poll

Prepared by Princeton Survey Research Associates International

September 2011

Results for the Massachusetts U.S. Senate Registered Voter Poll are based on telephone interviews with a random sample of 1,005 Massachusetts registered voters. Telephone interviews were conducted by landline (756) and cell phone (249, including 79 without a landline phone). The survey was conducted by Princeton Survey Research Associates International (PSRAI). Interviews were done in English by Braun Research, Inc. from September 22-28, 2011. Statistical results are weighted to correct known demographic discrepancies. The margin of sampling error for the complete set of weighted data is ± 3.8 percentage points. For multiple questions the sample was split randomly into two forms; 506 respondents were administered the Form A version of the interview, and sampling error for that form is ± 5.3 percentage points; 499 respondents were administered Form B, and sampling error for that form is ± 5.4 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 22-28, 2011. As many as seven 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.

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

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

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

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 $S_{LL} = 883$ and $S_{CP} = 288$. Additionally, we will estimate of the ratio of the size of landline sample frame to the cell phone sample frame $R = 0.99$.

The second stage of weighting balanced sample demographics to population parameters. The sample was balanced, by form, to match Massachusetts population parameters for sex, age, education, race, Hispanic origin, region, number of adults in household, and telephone usage. The basic weighting parameters came from a special analysis of the Census Bureau's 2010 Annual Social and Economic Supplement (ASEC) that included all households in the continental United States. The telephone usage parameter came from an analysis of 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 national population. Table 1 compares weighted and unweighted sample distributions to population parameters.

Table 1: Sample Demographics

	<u>Parameter</u>	<u>Unweighted</u>	<u>Weighted</u>
<u>Gender</u>			
	Male	48.8%	50.8%
	Female	51.2%	49.2%
<u>Age</u>			
	18-24	12.6%	6.2%
	25-34	15.7%	9.8%
	35-44	18.2%	16.2%
	45-54	20.5%	19.6%
	55-64	13.5%	17.3%
	65+	19.5%	25.9%
<u>Education</u>			
	Less than HS Grad.	10.5%	5.6%
	HS Grad.	33.0%	26.2%
	Some College	19.4%	20.8%
	College Grad.	37.1%	46.5%
<u>Race/Ethnicity</u>			
	White/not Hispanic	80.9%	83.4%
	Black/not Hispanic	5.5%	6.7%
	Hispanic	7.0%	4.9%
	Other/not Hispanic	6.6%	3.0%
<u>Household Phone Use</u>			
	LLO	7.6%	10.1%
	Dual	70.7%	81.1%
	CPO	21.8%	8.8%
<u>Region</u>			
	West/Central	23.0%	28.1%
	Southeast	12.0%	13.6%
	Outer suburbs	29.0%	33.1%
	Inner suburbs	23.0%	13.5%
	Boston proper	13.0%	11.7%
<u># of adults in HH</u>			
	One	16.7%	25.4%
	Two	52.3%	54.3%
	Three+	31.1%	20.3%

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 1.48.

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

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

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 ± 3.8 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 3.8 percentage points away from their true values in the population. The margin of error for results based on Form A respondents is ± 5.3 percentage points; for Form B respondents it is ± 5.4 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 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:²

- Contact rate – the proportion of working numbers where a request for interview was made³
- 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 24 percent. The response rate for the cellular samples was 21 percent.

² PSRAI's disposition codes and reporting are consistent with the American Association for Public Opinion Research standards.

³ 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

<u>Landline</u>	<u>Cell</u>	
7226	2384	Total Numbers Dialed
380	5	Non-residential
295	3	Computer/Fax
2		Cell phone
1858	416	Other not working
761	31	Additional projected not working
3930	1929	Working numbers
54.4%	80.9%	Working Rate
254	10	No Answer / Busy
938	659	Voice Mail
3	1	Other Non-Contact
2735	1259	Contacted numbers
69.6%	65.3%	Contact Rate
573	327	Callback
1172	484	Refusal
990	448	Cooperating numbers
36.2%	35.6%	Cooperation Rate
75	24	Language Barrier
0	105	Child's cell phone
127	39	Not a registered voter
788	280	Eligible numbers
79.6%	62.5%	Eligibility Rate
32	31	Break-off
756	249	Completes
95.9%	88.9%	Completion Rate
24.2%	20.7%	Response Rate