



I. SUMMARY

Abt SRBI conducted the 2013 Boston Mayoral Election Survey on behalf of the Center for Public Opinion at the University of Massachusetts Lowell. The survey included telephone interviews with a representative sample of 605 registered voters living in the city of Boston. Telephone interviews were conducted by landline (n=393) and cell phone (n=212, including 107 without a landline phone). Interviewing was conducted from October 2nd to 7th, 2013 in both English and Spanish. The margin of sampling error for weighted estimates based on registered voters is +/- 5 percentage points. Details on the sample design, weighting, and response rates are discussed below.

II. SAMPLE DESIGN

The target population for the survey is non-institutionalized persons age 18 and over living in the city of Boston who are registered to vote at their current address. Samples were drawn from both the landline and cellular random digit dial (RDD) frames to represent people with access to either a landline or cell phone. Both samples were provided by Survey Sampling International, LLC according to Abt SRBI specifications. Numbers for the landline sample were drawn with equal probability from active blocks (area code + exchange + two-digit block number) that contained one or more residential directory listings. The cell sample was drawn through a systematic sampling from 1000-blocks dedicated to cellular service.

In order to more efficiently reach cell phone respondents in this survey, Abt SRBI used information provided by the sampling vendor to target cell phone users who currently reside in the city of Boston: activity flag, the billing rate center associated with the number, and the zip code of the cell user's billing address, when available. The cell sample (n=29,000) was then stratified based on these three dimensions, as shown in Table 1. Activity flags indicated whether the number was estimated to be "active" or "inactive" based on the level of recent outgoing calling activity. The cell sample was also stratified based on whether the number could be matched to a billing zip code, and if so, whether that zip code was located in the city of Boston. About 49% of the numbers in the cell sample were able to be matched to a billing zip code. The cell sample was also stratified as "high density" or "low density" based on the proximity of each number's billing rate center to the city of Boston.



Table 1. Stratification of Cell Sample Purchased (n=29,000)

	Active		Inactive	
	High Density	Low Density	High Density	Low Density
Billing Zip Inside Boston	3291	496	367	92
Billing Zip Outside Boston	6617	2503	665	268
Billing Zip Not Available	7482	2367	3417	1435

To maximize the efficiency of reaching cell phone users in this survey, numbers classified as “inactive” or associated with billing zip codes outside of the city of Boston were excluded from the final sample. In addition, cell sample in the “low density” billing rate centers for which a billing zip code could not be appended were also excluded. The excluded cell sample is highlighted in red in Table 1. The remaining cases that were retained in the final cell sample fell into the following three strata: (1) numbers that were matched to a billing zip code in the city of Boston and associated a high density billing rate center, (2) numbers that were matched to a billing zip code in the city of Boston and associated a low density billing rate center, and (3) numbers that were not able to be matched to a billing zip code but were associated with a high density billing rate center. This stratification allowed us to control survey costs by increasing the amount of interviewer time spent dialing eligible numbers and minimizing the time spent dialing numbers for people who do not live in the study area and are therefore ineligible.

III. CALLING PROTOCOL

Interviews were conducted from October 2nd to 7th, 2013. As many as 5 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. For the landline sample, interviewers were 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. For the cell 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 voter were only asked demographic questions were weighting purposes.

IV. STATISTICAL WEIGHTING

The final weights produced for this survey compensated for the dual-frame sample design, including the stratification of the cell sample described above, and aligned the full sample (completed interviews plus non-registered voter screen-outs) to match the population parameters of the adult non-institutionalized population in the city of Boston. The weighting is based on the combined sample of 605 registered voters and 208 Boston residents who reported that they were not registered to vote (n=813).

First Stage Weighting

The first stage of weighting corrected for different probabilities of selection associated with selecting landline and telephone numbers. The weights were also adjusted for the overlapping landline and cell sample frames by adjusting to the estimated percentages of phone service (landline only, cell phone only or has both kinds of phones) and averaging the dual-users from the landline and cell phone together.

Second Stage Weighting

The second stage of weighting balanced sample demographics to estimated population parameters for the city of Boston. The full sample was balanced to match population parameters for sex, age, education level, race/Hispanic ethnicity, neighborhood, and telephone usage. The population parameters for sex, age, education, race, and Hispanic ethnicity were computed from the 2009-2011 American Community Survey (ACS), filtered on non-institutionalized adults aged 18 and older residing in the city of Boston. The population parameter for the respondent's neighborhood was obtained from the 2010 Census Summary File 1 by filtering on adults 18 and older living in Census tracts in each neighborhood. The telephone usage population estimates were constructed from the model-based estimates for Suffolk County, Massachusetts, that were released by the National Center for Health Statistics for the year 2011¹. Since the cell phone-only adult population has

¹ Blumberg, Stephen J. and Julian V. Luke. Wireless Substitution: State-level Estimates From the National Health Interview Survey, 2010-2011. National Center for Health Statistics. October 2012.

increased every year since 2011, these sub-state level estimates were updated to reflect national trends according to the 2012 NCHS report.²

The second-stage weighting was conducted using an operation known as raking ratio estimation, or “raking”. Raking is used to reduce the risk of biases due to nonresponse and non-coverage in sample surveys. The raking procedure uses an iterative technique that simultaneously calibrates the sample to population distributions defined by socio-demographic parameters. After the raked weights were generated, we examined the distribution of values. Weights were trimmed at 0.20 and 3.57 to prevent individual interviews (i.e., those with large weights) from having too much influence on the final results. This trimming process also served to reduce the variance of the weight values, and, in turn, reduce the design effect from weighting. The use of these weights in statistical analysis ensures that the demographic characteristics of the full sample closely approximate the demographic characteristics of the adult population in the city of Boston. In the survey dataset, this full sample weight is labeled *WEIGHT*. Table 2 compares weighted and unweighted total sample distributions to population parameters.

Table 2. Demographics of the Full Sample (n=813)

	Benchmark	Weighted	Unweighted
Male	46.4%	47.2%	47.9%
Female	52.6%	52.8%	52.2%
18-29	37.6%	35.8%	18.0%
30-39	17.7%	18.1%	14.3%
40-49	14.3%	14.6%	14.9%
50-64	17.8%	18.2%	26.8%
65+	11.6%	12.1%	23.3%
DK/REF	1.0%	1.2%	2.8%
Less Than High School	13.3%	13.4%	10.5%
High School Graduate	21.3%	21.4%	21.5%
Some College	26.3%	25.4%	19.7%
College Graduate	23.1%	23.2%	24.0%
Post-graduate	15.0%	15.4%	21.2%

² Blumberg, Stephen J. and Julian V. Luke. Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July – December 2012.



DK/REF	1.0%	1.3%	3.2%
White only Non-Hispanic	51.5%	50.3%	48.1%
Black only Non-Hispanic	20.3%	20.5%	19.4%
Hispanic (any race)	14.9%	15.4%	17.2%
Other race/Multi-Race Non-Hispanic	12.4%	12.6%	11.2%
DK/REF	1.0%	1.3%	4.1%
Allston/Brighton	13.4%	13.0%	10.6%
Charlestown, South Boston, South End, East Boston	18.7%	18.7%	20.2%
Downtown: North End/Downtown, Beacon Hill, Back Bay, Fenway, Chinatown	17.2%	16.7%	13.8%
Dorchester, Roxbury, Mattapan, Hyde Park	31.6%	32.3%	31.6%
Roslindale, Jamaica Plain, West Roxbury	18.1%	18.1%	19.7%
DK/REF	1.0%	1.2%	4.2%
Cell Phone Only	39.4%	38.0%	24.6%
Dual User	49.4%	50.3%	58.2%
Landline Only	10.2%	10.4%	12.2%
DK/REF	1.0%	1.4%	5.0%

V. DESIGN EFFECT AND MARGIN OF ERROR

Weighting and survey design features that depart from simple random sampling tend to result in an increase in the variance of survey estimates. This increase, known as the design effect or *deff*, should be incorporated into the margin of error, standard errors, and tests of statistical significance. The design effect is computed as the variance in the weights (w_i) in relation to the expected variance from an otherwise simple random sample:

$$deff = n \times \frac{\sum w_i^2}{\sum w_i}$$

The design effect for the sample of registered voters (n=605) is 1.58. The margin of error for an estimated percentage of 50% for the complete set of weighted data is approximately +/- 5 percentage points at the 95% level of confidence. This means that in 95 out of every 100 samples



drawn using the same methodology, estimated proportions based on the sample of registered voters will be no more than 5 percentage points away from their true values in the population. It is important to remember that random sampling error is only one possible source of error in a survey estimate. Other sources, such as question wording and reporting inaccuracy, may contribute additional error.

VI. RESPONSE RATES

Table 3 reports the disposition of all sampled telephone numbers dialed for the survey. Abt SRBI calculates four component rates: Response rate, Cooperation rate, Refusal rate and Contact rate³:

- Response rate – the number of complete interviews with reporting units divided by the number of eligible reporting units in the sample.
- Cooperation rate – the proportion of all cases interviewed of all eligible units ever contacted.
- Refusal rate – the proportion of all cases in which a housing unit or respondent refuses to do an interview
- Contact rate – measures the proportion of all cases in which some responsible member of a housing unit was reached by the survey

The response rate for the landline sample ranged from 5.8 to 18.1%. The response rate for the cellular sample ranged from 14.9 to 18.1%.

Table 3. Sample Dispositions

		Landline	Cell
Interview (Category 1)			
Complete	1.000	393	212
Screen-outs	1.100	137	878
Partial	1.200	14	5
Eligible, non-interview (Category 2)			
Refusal and breakoff	2.100	22	8
Refusal	2.110	1085	1407
Non-contact	2.200	0	0

³ Abt SRBI's disposition codes and reporting are consistent with the American Association for Public Opinion Research standards.



Respondent never available	2.210	22	58
Answering machine household-no message left	2.221	402	254
Physically or mentally unable/incompetent	2.320	77	53
Household-level language problem	2.331	240	235
Unknown eligibility, non-interview (Category 3)			
Always busy	3.120	423	371
No answer	3.130	3895	722
Call blocking	3.150	1	7
Technical phone problems	3.160	0	0
No screener completed	3.210	598	839
Not eligible (Category 4)			
Fax/data line	4.200	981	5
Non-working/disconnect	4.300	21047	820
Temporarily out of service	4.330	160	243
Cell phone	4.420	6	3
Business, government office, other organizations	4.510	1634	272
Other	4.900	4	0
Total phone numbers used		32956	8667
Completes and Screen-Outs (1.0/1.1)	I	530	1090
Partial Interviews (1.2)	P	14	5
Refusal and break off (2.1)	R	1107	1415
Non Contact (2.2)	NC	424	312
Other (2.3)	O	317	288
Unknown household (3.1)	UH	6134	3375
Unknown other (3.2, 3.9)	UO	598	839
Not Eligible (4.0)	NE	23832	1343



e = Estimated proportion of cases of unknown eligibility that are eligible.	$(I+P+R+NC+O)/((I+P+R+NC+O)+N)$	0.091	0.698
Response Rate 1	$I/(I+P) + (R+NC+O) + (UH+UO)$	0.058	0.149
Response Rate 2	$(I+P)/(I+P) + (R+NC+O) + (UH+UO)$	0.060	0.150
Response Rate 3	$I/((I+P) + (R+NC+O) + e(UH+UO))$	0.176	0.180
Response Rate 4	$(I+P)/((I+P) + (R+NC+O) + e(UH+UO))$	0.181	0.181
Cooperation Rate 1	$I/(I+P)+R+O$	0.269	0.390
Cooperation Rate 2	$(I+P)/((I+P)+R+O)$	0.276	0.391
Cooperation Rate 3	$I/((I+P)+R)$	0.321	0.434
Cooperation Rate 4	$(I+P)/((I+P)+R)$	0.329	0.436
Refusal Rate 1	$R/((I+P)+(R+NC+O) + UH + UO)$	0.121	0.193
Refusal Rate 2	$R/((I+P)+(R+NC+O) + e(UH + UO))$	0.368	0.234
Refusal Rate 3	$R/((I+P)+(R+NC+O))$	0.463	0.455
Contact Rate 1	$(I+P)+R+O / (I+P)+R+O+NC+ (UH + UO)$	0.216	0.382
Contact Rate 2	$(I+P)+R+O / (I+P)+R+O+NC + e(UH+UO)$	0.655	0.462
Contact Rate 3	$(I+P)+R+O / (I+P)+R+O+NC$	0.823	0.900