



MPR News / APM Research Lab

Ground Level Survey of Minnesotans
October 2017

Methodology report prepared by
Robyn Rapoport, AJ Jennings and David Yuchimiuk
1 Braxton Way
Suite 125
Glen Mills, PA 19342





Contents

OVERVIEW	3
SAMPLE DESIGN	3
RDD Sample Generation.....	3
Smart Cell Sample Generation.....	4
SSRS Omnibus Prescreened Sample.....	4
Household and Respondent Selection	4
DATA COLLECTION	5
Programming and Testing.....	5
Pretesting.....	5
Training Materials and Interviewer Training.....	5
Call Rules and Refusal Avoidance and Conversion Strategies	5
Field Period and Weekly Updates	6
Data Processing.....	6
Final Data.....	6
WEIGHTING.....	8
Design Weight:.....	8
Post stratification weighting:.....	9
Weight truncation ('trimming'):.....	10
Design Effect and Margin of Sampling Error	12
RESPONSE RATES.....	13
DELIVERABLES.....	13



OVERVIEW

Minnesota Public Radio is one of the nation's premier public radio stations. The 45-station radio network serves nearly all of Minnesota and parts of the surrounding states and reaches one million listeners each week. The APM Research Lab is a division of American Public Media established to help bring credible research and data to the broader public. The goal of the MPR News / APM Research Lab Ground Level Survey is to seek the opinions and attitudes of adult Minnesotans on a variety of topics including financial security and related issues and institutions in Minnesota, such as the police, public schools and the state government.

MPR | APM contracted with SSRS to manage data collection of the 2017 Ground Level Survey. The survey was conducted via landline and mobile telephone with a representative sample of Minnesota respondents, age 18 and older, including an oversample of Minnesotans of color and over/undersampling by geography to obtain a minimum of 200 interviews in each of the eight client-designated strata. The geographic strata include: Minneapolis/St. Paul; 7-county region minus M/SP (Twin Cities suburbs); the cities of Duluth, St. Cloud, and Rochester, including those living in the areas immediately surrounding those three cities; and three multi-county rural regions of the state: Northern MN, Central MN, and Southern MN (excluding the three aforementioned cities). A total of 1,654 interviews were completed from August 22 through September 14, 2017. Approximately, forty-two percent of the interviews were conducted on landlines (n=695) and fifty-eight percent of the interviews were conducted on cell phones (n=959).

This report is organized into six sections. The first section provides an overview of the project. The next two sections describe the sample design and data collection and fielding. The final three sections address weighting procedures, response rates, and project deliverables.

SAMPLE DESIGN

The sample plan was designed to obtain 200 interviews in each of the following geographic areas/counties: Minneapolis/St. Paul, 7-county region minus M/SP (Twin Cities suburbs); the cities of Duluth, St. Cloud, and Rochester, including those living in the areas immediately surrounding those three cities; Northern MN (minus Duluth), Central MN (minus St. Cloud), and Southern MN (minus Rochester).¹ In order to attain a minimum of 200 interviews in each strata and maximize the incidence of non-white interviews (e.g., African Americans, Hispanics, American Indians, and Asians), landline and mobile RDD sample was blended with prescreened and geographically targeted (SSI Smart Cell)² samples. The RDD Landline and Cell phone sample was stratified by geography to over/undersample strata, as needed, to obtain the minimum of 200 interviews across all sample types.

RDD Sample Generation

The RDD landline and cell phone sample was generated by MSG's GENESYS sample generation system. Twin Cities, Duluth, St. Cloud, and Rochester were pulled based on zip code and converted into exchanges and RDD landline sample was pulled. For Twin City suburbs, North, Central and South, sample was drawn based on county-based definitions and the exchanges from Twin Cities, Duluth, St. Cloud, and Rochester

¹ Please contact the APM Research Lab (www.apmresearchlab.org) for the specific ZIP codes boundaries used for the Duluth, St. Cloud, and Rochester areas, and the specific counties included in each of the rural regions.

² <https://www.surveysampling.com/about/news/2016/ssi-launches-smart-cell-sample-increasing-incidence-rates/>



were excluded from these strata. The same process was followed for mobile phone generation, except that instead of exchanges, rate centers were used. Following sample generation, the landline sample was prepared using MSG's proprietary ID+ procedures, which limits sample to working phone banks, and also identifies and eliminates most of the non-working and business numbers and ported cell phone numbers. The RDD cell phone sample was prepared using MSG's Cell-WINS screening process to remove inactive numbers within the cellular RDD sample.

Smart Cell Sample Generation

Using RDD cell phone sample to target small geographic populations is particularly challenging since rate centers do not closely align with geographic boundaries. Given these limitations, SSRS obtained Smart Cell Sample from SSI for the three lowest-incidence strata, Duluth, St. Cloud, and Rochester in order to supplement the completions on cell phones for these strata. SSI Smart Cell uses databased sources that link addresses and cell phone numbers to be able to more precisely target geography than can be done with rate center selection alone.

SSRS Omnibus Prescreened Sample

The prescreened sample was obtained from screening respondents who were age 18+ and lived in Minnesota, through the SSRS Omnibus from January 2014 to August 2017. Since it is extremely challenging to oversample non-whites in Minnesota based on geography, particularly given the fact that non-whites only modestly cluster within the Twin Cities, non-whites were oversampled within the prescreened sample by simply increasing the sampling portion on non-white prescreened sample in relation to white prescreened sample.

TABLE 1: Population Counts

Strata	Strata Name	Total Population	Share of MN Pop	Population of Color age 18+	Share of strata population	Total Interviews
1	MSP/Central Cities	694,993	13%	191,428	35%	207
2	Twin Cities Suburbs	2,257,121	42%	283,464	17%	211
3	Duluth area	117,077	2%	7,485	8%	201
4	St. Cloud area	126,764	2%	10,862	11%	200
5	Rochester area	121,737	2%	16,016	17%	218
6	Rural Northern MN	379,226	7%	25,290	9%	214
7	Rural Central MN	834,841	15%	35,221	6%	201
8	Rural Southern MN	887,412	16%	56,216	8%	202
TOTAL		5,419,171				1,654

Household and Respondent Selection

In the RDD and prescreened landline frames, a within-household selection procedure was utilized in order to reduce potential for bias created when the person responding to the survey is the one more likely to answer the phone. Specifically, the adult who answered (or was brought to) the phone in the landline frame was asked how many adults age 18 or older live in the household. Households with more than one adult in this age range were randomly assigned to ask for the youngest adult male or female currently at home 75% of the time and the oldest adult male or female currently at home the other 25% of the time. If no males/females



were at home, the instrument then asked for the youngest adult of the opposite gender (i.e., female/male) currently at home 75% of the time and the oldest adult of the opposite gender 25% of the time. Cell phones are considered individual devices rather than belonging to a household, and therefore the person answering the cell phone was the one who was interviewed if he or she was 18 years of age or older and their main residence was in Minnesota.

DATA COLLECTION

The instrument was designed by MPR | APM, while SSRS reviewed and offered suggestions for possible edits. Specifically, SSRS provided feedback regarding question wording, order, clarity, and other issues pertaining to questionnaire quality. Prior to the field period, SSRS formatted the questionnaire and the study was programmed into SSRS's Computer Assisted Telephone Interviewing (CATI) system. SSRS conducted the pretest of the instrument on August 15, 2017. Interviews were conducted from August 22 through September 14, 2017.

Programming and Testing

Prior to the field period extensive checking of the program was conducted to assure that skip patterns followed the design of the questionnaire. Also, thousands of cases of random data were run through the program to provide another check of the accuracy of the CATI program. The project team reviewed all random data for accuracy.

Pretesting

SSRS conducted the pretest on August 15 to achieve a total of 19 pretest completes. For the pretest interviews, SSRS used listed landline and cellphone sample. SSRS provided MPR | APM with a detailed pretest memo including feedback on the overall instrument and interviewing quality.

Training Materials and Interviewer Training

Prior to the start of the study, interviewers received both written materials on the survey and formal training for conducting the survey. SSRS's project team briefed and trained interviewers on the issues specific to the study, explaining the study's overall objectives, specific procedures, and questionnaire content.

Interviewer training was conducted prior to the pretest and immediately before the survey was officially launched. Call center supervisors and interviewers were walked through each question in the questionnaire. Interviewers were given instructions to help them maximize response rates and ensure accurate data collection. They were instructed to encourage participation by emphasizing the importance of the project and to reassure respondents that the information they provided was confidential. Interviewers were monitored during the first several nights of interviewing and provided feedback, where appropriate, to improve interviewer technique and clarify survey questions. The project team monitored interviews throughout the field period and provided feedback, when necessary, to ensure that best practices were being followed.

Call Rules and Refusal Avoidance and Conversion Strategies

SSRS carried out several strategies to maximize survey response by minimizing non-response and maximizing refusal conversion. The survey fielding enacted the following best-practice procedures:

- The call rule included one initial call plus an average of five callbacks.



- To increase the probability of completing an interview, a differential call rule was established that required that call attempts be initiated at different times of day and different days of the week.
- Landline sample was power-dialed, using a computer to dial the number, to reduce dialing errors. Cell phone sample was dialed manually in keeping with TCPA regulations.
- Respondents were permitted to schedule call-back times.

Field Period and Weekly Updates

Interviews were conducted from August 22 through September 14, 2017. Throughout the field period, SSRS provided MPR | APM with weekly updates with key information that tracked overall progress of the study. These reports, designed to provide snapshot information of key variables of interest, included tables for completes per sample type by gender, age, ethnicity, strata, PartyID/Lean, personal finances, MN future, Trump approval, and voter results in the last election. Along with the weekly data reports, SSRS provided a narrative regarding field progress and reported on any field-related concerns.

Data Processing

Data were checked using multiple methods. Data were checked after the first night of interviewing and throughout the field period to confirm that skip patterns were correctly followed. In addition, the back-end programmer created a program consisting of instructions derived from the skip patterns designated on the questionnaire to check the data. The program confirmed that data were consistent with the definitions of codes and ranges and matched the appropriate bases of all questions. The SSRS team also reviewed preliminary SPSS files and conducted an independent checking of all created variables to ensure that all variables were accurately constructed.

Final Data

Table 2 shows final counts by gender, age, ethnicity, and strata.



TABLE 2: Final Counts by Gender, Age, Ethnicity & Strata

Gender	LL	LL %	Cell	Cell%	Pre-screened	Pre-screened%	Total	Total %
Gender								
Male	166	43%	372	57%	309	50%	847	50%
Female	212	55%	272	42%	309	50%	793	48%
Other	2	1%	5	1%	0	0%	7	0%
Refused	2	1%	2	0%	3	0%	7	0%
Age								
18 to 24	4	1%	66	10%	27	4%	97	6%
25 to 34	7	2%	102	16%	57	9%	166	10%
35 to 44	26	7%	95	15%	51	8%	172	10%
45 to 54	56	15%	103	16%	91	15%	250	15%
55 to 64	83	22%	132	20%	127	20%	342	21%
65 to 74	101	26%	97	15%	126	20%	324	20%
75+	97	25%	49	8%	140	23%	286	17%
65+ (unspecified)	1	0%	0	0%	0	0%	1	0%
Refused	7	2%	7	1%	2	0%	16	1%
Race								
White non-Hispanic	343	90%	559	86%	497	80%	1399	85%
Of color	34	9%	73	11%	107	17%	214	13%
Refused	5	1%	19	3%	17	3%	41	2%
Region/Strata								
MSP/Central Cities	28	7%	34	5%	145	23%	207	13%
TC Suburbs	22	6%	85	13%	104	17%	211	13%
Duluth	50	13%	105	16%	46	7%	201	12%
St. Cloud	69	18%	92	14%	39	6%	200	12%
Rochester	83	22%	100	15%	35	6%	218	13%
Northern MN	51	13%	85	13%	78	13%	214	13%
Central MN	41	11%	69	11%	91	15%	201	12%
Southern MN	38	10%	81	12%	83	13%	202	12%
Total	382	100%	651	100%	621	100%	1654	100%



WEIGHTING

Survey data were weighted to: 1) adjust for the fact that not all survey respondents were selected with the same probability and 2) account for gaps in coverage in the survey frame.

Design Weight:

In the first stage, SSRS developed design weights to compensate for a range of known biases that occur in telephone interviewing in general and the sample design for this study specifically. These are summarized below:

- **TOTAL PROBABILITY OF SELECTION WEIGHT=** The weighting process takes into account the disproportionate probabilities of household and respondent selection due to the number of separate telephone landlines and cellphones answered by respondents and their households, as well as the probability associated with the random selection of an individual household member.

Probability of Selection (Pphone): A phone number's probability of selection depends on the number of phone-numbers selected out of the total sample frame. So for each respondent whose household has a landline phone number, this is calculated as total landline numbers dialed divided by total numbers in the landline frame and conversely for respondents answering at least one cell phone number, this is calculated as total cell phone numbers divided by total numbers in the cell phone frame.

Probability of Respondent selection (Pselect): In households reached by landline, a single respondent is selected. Thus, the probability of selection within a household is inversely related to the number of adults in the household.

Total Probability of Selection: This is calculated as the phone number's probability of selection (by frame), multiplied by the number of devices of each type the respondent answers, and for landlines, divided by the number of adults in the household.³ Thus, for each respondent a probability can be calculated for being reached via landline (LLprob) and for being reached via cell phone (Cellprob). These calculations are:

$$\begin{aligned} \text{LLprob} &= \text{Pphone} * \text{Pselect} \\ \text{Cellprob} &= \text{Pphone} \end{aligned}$$

The sample weights derived at this stage are calculated as the inverse of the combined probability of selection, or:

$$1/(\text{LLprob} + \text{Cellprob} - \text{LLprob} * \text{CellProb})$$

- **PROPENSITY CORRECTION WEIGHT=** This adjustment accounts for the potential bias associated with re-contacting respondents through the use of pre-screened SSRS Omnibus sample. Inverse Probability Weighting (IPW), or Propensity Weighting, is typically used to adjust for attrition in

³ To avoid extremely large or small weights, the maximum number of devices for each type of phone, and the maximum number of adults was capped at 3.



longitudinal studies. Characteristics of the respondents as measured in the initial study are used to model the respondents' propensity to respond to the re-contact survey.

The predictive values from the logistic regression model were used as the probability of a person completing the re-contact survey. The inverse of this probability was used as the propensity weight. As such, higher propensity weights correspond to respondents who have a lower probability of responding to the re-contact survey.

- LISTED CELLPHONE WEIGHT = The total listed cellphone sample was balanced back to the true distribution of listed cell sample in the RDD frame.
- NON-WHITE OVERSAMPLING WEIGHT = The proportion of dialed Non-White prescreened sample was weighted to the overall Non-White prescreened sample available to correct for disproportionate dialing of this group.
- STRATIFICATION WEIGHT = Rebalancing the completed interviews distribution to the population counts.
- DESIGN WEIGHT = Is the product of each of the design weight adjustments (Probability of Selection * Propensity Correction * Listed Cellphone * Non-White Oversampling * Stratification Weight)

Post stratification weighting⁴:

With the design weight applied, the sample underwent the process of iterative proportional fitting (IPF), in which the sample was balanced to match known adult-population parameters based on U.S. Census Bureau's 2011-2015 American Community Survey, and official election results data. These data were assembled by SSRS in collaboration with the APM Research Lab. This process of weighting was repeated until the root mean square error for the differences between the sample and the population parameters was 0 or near-zero.

The post-stratification stage consisted of a two-step raking procedure. The first raking procedure occurred at the strata level, where targets were set for age (less than age 45, and age 45+), education⁵ (High School or Less and Post-High School), and share of votes in the 2016 presidential elections⁶. A second statewide raking procedure was conducted where the population parameters were: age (18-34; 35-54; 55-64; 65+), gender, race/ethnicity (white vs non-white); education⁷ (high school graduate or less, some college or two-

⁴ Missing data were imputed using a Hot Deck procedure prior to raking.

⁵ Respondents age 18-24 were assigned into its own category for the per-region education weight as the benchmarks for education corresponded to age 25+ data from the ACS 2011-2015.

⁶ Persons who refused or said they did not know who they voted for were imputed using a standard Hot Deck imputation procedure based on party identification. <http://www.sos.state.mn.us/elections-voting/2016-general-election-results/> was the source for Duluth, Rochester and St. Cloud strata at the precinct-level; <http://www.sos.state.mn.us/elections-voting/2016-general-election-results/2016-precinct-results-spreadsheet/>, included the precincts that fell within the zip codes included in those areas.

⁷ The 2011-2015 ACS data for education exclude the population under age 25. To account for this, at the overall second-stage rake, the educational attainment of the population under the age of 25 was estimated based on that of this population in Minnesota overall (using the 2015 ACS), and added to the ACS 2011-2015 estimates for the population 25 and older.



year degree, four-year college, post-graduate degree or more), phone-usage ⁸ (cell phone only vs. other), Minnesota region and share of votes in the 2016 presidential elections.

Weight truncation ('trimming'):

To reduce variance caused by extremely large weights, the weights were truncated to top/bottom 2%. Data have an overall margin of error of +/- 4.2% at the 95 percent confidence level.

Weights were normalized such that the sum of weights equals the un-weighted number of completed interviews. SSRS also provided MPR | APM a second weight normalized to the adult population of Minnesota.

Table 3: Regional Weighting Parameters

	Minneapolis /St Paul	Seven-county region minus M/SP (Twin Cities Suburbs)	Duluth	St. Cloud	Rochester	Northern MN	Central MN	Southern MN
Education								
Age 18-24	17.10%	9.70%	20.50%	20.00%	10.50%	9.80%	9.90%	13.40%
High school or Less	24.90%	22.50%	25.60%	25.00%	21.70%	33.70%	36.40%	34.60%
Post-High school	58.10%	67.80%	53.80%	55.00%	67.80%	56.60%	53.80%	52.00%
Age								
Less than age 45	59.60%	46.90%	48.70%	56.40%	48.20%	37.20%	40.90%	43.30%
Age 45+	40.40%	53.10%	51.30%	43.60%	51.80%	62.80%	59.10%	56.70%
Share of 2016 Votes								
Clinton	77.8%	49.0%	56.4%	37.9%	48.9%	38.4%	29.1%	34.6%
Trump	13.7%	41.9%	34.2%	53.0%	40.7%	53.6%	63.3%	56.9%
Other	8.4%	9.0%	9.4%	9.1%	10.3%	8.1%	7.6%	8.5%

⁸ Phone status was obtained from the NHIS Modeled estimates (with standard errors) of the percent distribution of household telephone status for adults aged 18 and over, by state: United States, 2015, and projected for year 2017 for Cellphone Only (CPO) vs. Not-CPO.



Table 4: Statewide Weighting Parameters

Minnesota	
Age	
18-34	29.7%
35-54	34.7%
55-64	17.1%
65+	18.6%
Gender	
Male	49.6%
Female	50.4%
Race	
White	84.9%
Non-White	15.1%
Education	
High School or Less	34.1%
Some College	33.7%
College	22.1%
Post-graduate degree	10.1%
Phone Status	
Cellphone Only	53.2%
Not-Cellphone Only	46.8%
Region	
Minneapolis/St. Paul	13.0%
Twin Cities Suburbs (seven-county region minus M/SP)	41.2%
Duluth	2.3%
St. Cloud	2.4%
Rochester	2.2%
Northern MN (Northland and Northwest minus Duluth)	7.1%
Central MN (Central and West Central minus St. Cloud)	15.2%
Southern MN (Southern and Southwestern minus Rochester)	16.5%
2016 Share of Votes	
Clinton	46.4%
Trump	44.9%
Other	8.7%



Design Effect and Margin of Sampling Error

Weighting procedures increase the variance in the data, with larger weights causing greater variance. Complex survey designs and post-data collection statistical adjustments affect variance estimates and, as a result, tests of significance and confidence intervals. Table 5 reports the weight-adjusted margins-of-error overall and by strata.

TABLE 5: Design Effect & Margins of Error

	Sample Size	Design Effect	% Margin of Sampling Error 95% confidence level (+/-)	% Margin of Sampling Error 90% confidence level (+/-)
Total	1654	3.0	4.2	3.5
MSP/Central Cities	207	2.1	9.8	8.3
TC Suburbs	211	1.4	8.1	6.8
Duluth area	201	2.5	10.8	9.1
St. Cloud area	200	3.0	12.0	10.1
Rochester area	218	1.7	8.7	7.3
Rural Northern MN	214	2.1	9.7	8.2
Rural Central MN	201	1.7	9.0	7.6
Rural Southern MN	202	1.8	9.4	7.9



RESPONSE RATES

The response rates for this study (shown in Table 6) were calculated using AAPOR's RR3.

First, we calculated the response rate for each sample type: main landline, main Cell, and prescreened sample (Table 6). We then multiplied the response rate from the prescreened sample by the average SSRS Omnibus response rate (7%). The combined response rates and the overall response rate are shown in the last row of Table 6.

TABLE 6: Response Rates

Disposition	Main LL	Main CELL	Prescreened Total	Total
Eligible, Interview (Category 1)				
Complete	382	651	621	1,654
Eligible, non-interview (Category 2)				
Refusal and breakoff	0	80	9	89
Break off	186	492	95	773
Answering machine household-message left	926	3,573	279	4,778
Physically or mentally unable/incompetent	11	20	4	35
Language problem	11	47	1	59
Unknown eligibility, non-interview (Category 3)				
Always busy	448	425	11	884
No answer	3,058	3,171	301	6,530
Call blocking	360	245	41	646
Housing unit, unknown if eligible respondent	901	1,653	134	2,688
No screener completed	75	1,065	154	1,294
Not eligible (Category 4)				
Fax/data line	358	96	9	463
Non-working number	20,650	5,338	194	26,182
Business, government office, other organizations	366	299	28	693
No eligible respondent	10	399	38	447
Total phone numbers used	27,742	17,733	2,051	47,526
MPR Response Rate	21.8%	14.5%	39.2%	
MPR Response Rate + SSRS Omnibus Response Rate	21.8%	14.5%	2.8%	11.8%

DELIVERABLES

SSRS delivered to MPR | AMP (1) final weighted SPSS dataset, (2) final weighted banners in Microsoft Word format, (3) final methodology report and (4) final reporter contact dataset.