



TELEPHONE CARAVAN® Methodology

The following pages describe the methodology used for the ORC International Telephone CARAVAN® survey conducted January 29-February 1, 2015.

The study was conducted using two probability samples: randomly selected landline telephone numbers and randomly selected mobile (cell) telephone numbers. The combined sample consists of 1,009 adults (18 years old and older) living in the continental United States. Of the 1,009 interviews, 509 were from the landline sample and 500 from the cell phone sample. The margin of error for the sample of 1,009 is +/- 3.09 at the 95% confidence level. Smaller subgroups will have larger error margins.

Surveys are collected by trained and supervised US based interviewers using ORC International's computer assisted telephone interviewing (CATI) system. Final data is adjusted to consider the two sample frames and then weighted by age, gender, region, race/ethnicity and education to be proportionally representative of the US adult population.

As a founding member of the Code of Standards of the Council of American Survey Research Organizations (CASRO) and a member of the European Society for Opinion and Marketing Research (ESOMAR), we adhere to a rigorous Code of Standards and Ethics for Survey Research. As required by CASRO, we will maintain the anonymity of our respondents. No information will be released that in any way will reveal the identity of a respondent. Our authorization is required for any publication of the research findings or their implications.

Sampling

Telephone CARAVAN® uses a dual frame sampling design. This means that the sample is drawn from two independent sample frames—one for landlines and one for cell phones.

Landline Sample

ORC International's Random Digit Dial (RDD) landline telephone sample is generated using the following process. First, all residential exchanges and working 100 banks are determined. A 100 series bank is defined to be working if one (1) or more assigned residential landline telephone numbers are found within that bank. Within any given 100-series bank there are 100 possible two-digit combinations that form the suffix in a complete telephone number. For instance, in working bank 21, numbers 00-99 can be appended to form the one



hundred complete numbers 2100-2199. This forms the sample frame of all possible telephone numbers within which RDD samples are then generated.

All exchanges and associated working 100 series banks within the RDD database are arrayed in a pre-defined order – specifically, exchanges are ordered within predominant county based on HH size largest to smallest. Counties are ordered by Metro/non-Metro areas within State within Census Division. Using the ordered database, an **Epssem** (Equal Probability Selection Method) sample is generated. A sample frame is defined to be a set of 100 series banks within exchanges serving a particular geographic area (city, county, state, National, etc.). The sampling interval is determined by dividing the total number of possible numbers in the geographic sample frame by the number of requested RDD sample records. Each sample interval is of equal size. One number is selected at random from each sample interval. All possible telephone numbers are given an equal probability of selection regardless of the density of assigned residential households within them.

Cell Phone Sample

The MSG Cellular RDD database is constructed quarterly utilizing Telecordia's LERG product. The LERG is a continuously updated suite of telephony databases that, among other things, provides current information for every active Thousand Series Block in the North American Numbering Plan. Using multiple files within the LERG, every thousand series block that is dedicated to providing wireless service is identified and incorporated into the Cellular RDD database. Additionally, Rate Center boundaries are constructed each quarter. Rate Center boundaries provide the information needed to map each cellular thousand series block to a predominant county.

The cell phone sample was generated from cell phone 1,000 series blocks with all the 100 series banks within each block turned on. The sampling interval is then calculated by dividing the universe of all possible numbers by the number of records desired, thus specifying the size of the frame subdivisions. Within each of the subsets one number is selected at random giving all numbers an equal probability of selection.

At this point, the frame size has been fixed and divided into equal-sized subsets of ten-digit numbers. Within each of the subsets one number is selected at random giving all numbers an equal probability of selection.

Weighting

In probability-based samples such as CARAVAN®, the basis of the weighting is the inverse of the selection probability. Then, weighting adjustments are frequently used to reduce the potential for biases that may be present due to incomplete frame coverage and

survey nonresponse--both inherent in all telephone surveys. These adjustments may take advantage of geographic, demographic, and socioeconomic information that are known for the population and measured in the sample surveys. The adjustments reduce potential bias to the extent that the survey respondents and nonrespondents (noncontacts, refusals, etc.) with similar geographic, demographic, and socioeconomic characteristics are also similar with respect to the survey statistics of interest. In other words, post-survey weighting adjustments reduce bias if the weighting variables are related to (correlated with) the survey measures and the likelihood of survey participation.

The CARAVAN® *landline-cell* combined sample is a dual frame sampling design. This means that the sample is drawn from two independent sampling frames—one for landlines and one for cell phones. Adults with a landline but no cell phone (A) must be reached through a landline telephone sample. Adults with a cell phone and no landline (C) must be reached through the cell phone sample. Adults with both a landline and a cell phone (B) can be reached through either of the frames. Sampling from the two frames results in these four groups:

- a_1 : Landline respondents without a cell phone (landline only)
- b_1 : Landline respondents with a cell phone (dual user)
- b_2 : Cell phone respondents with a landline (dual user)
- c_2 : Cell phone respondents without a landline (cell only)

The dual user groups (b_1, b_2) are further classified into two subgroups:

Cell mostly: those who receive most calls on a cell phone

Landline mostly/Mixed use: those who receive most calls on a landline or who receive calls on both regularly

The National Health Interview Survey (NHIS) provides estimates of these user group populations. We weight-adjust the landline sample and the cell sample to their respective population proportions as reported from the NHIS. Once this design weight is calculated, the combined sample is weighted to represent the US population using data from the US Census Bureau's Current Population Survey (CPS). This form of weighting is referred to as *calibration weighting*ⁱ in that survey respondents are assigned weights that are calibrated to reflect the population. The calibration weighting for CARAVAN® is based on a series of ratio adjustments called iterative proportional fitting, or "*raking*"ⁱⁱ, which was first introduced by Deming and Stephan for use in the 1940 US census.



Definition of Classification Terms

The following definitions are provided for some of the standard demographics by which the results are tabulated. Other demographics are self-explanatory.

Income

The income groupings refer to the total household income before taxes.

Geographic Region

The states are contained in four geographic regions as follows:

North East

- New England: Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut
- Middle Atlantic: New York, New Jersey, Pennsylvania

Midwest

- East North Central: Ohio, Indiana, Illinois, Michigan, Wisconsin
- West North Central: Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas

South

- South Atlantic: Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida
- East South Central: Kentucky, Tennessee, Alabama, Mississippi
- West South Central: Arkansas, Louisiana, Oklahoma, Texas

West

- Mountain: Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada
- Pacific: Washington, Oregon, California, Hawaii, Alaska



About ORC International

ORC International is a collaborative and consultative research partner to hundreds of organizations around the globe. We possess a wide variety of resources, tools and technologies to collect and analyze information for our clients.

ORC International is ISO 20252 certified. To achieve certification, ORC International passed a comprehensive, on-site audit. The certification establishes globally recognized terms, definitions, and service requirements for project management in research organizations. Processes outlined in ISO 20252 are designed to produce transparent, consistent, well documented and error-free methods of conducting and managing research projects. Adherence and certification to such standards provides a basis of confidence for clients and other constituencies that the work produced is being executed with quality processes and controls in place. The internationally recognized standard also provides a basis for subcontractor evaluation.

ⁱ For a summary of calibration weighting, refer to Kalton, G. and I. Flores-Cervantes (2003) “Weighting Methods”, *Journal of Official Statistics*.

ⁱⁱ Deming, W. E. and F. F. Stephan (1940) “On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known,” *Annals of Mathematical Statistics*.