

CARAVAN® Landline/Cell Phone Methodology

This report presents the findings of a telephone survey conducted among two national probability samples, which, when combined, consists of 1,008 adults, 503 men and 505 women 18 years of age and older, living in the continental United States. Interviewing for this CARAVAN® Survey was completed on February 7-10, 2013. 654 interviews were from the landline sample and 354 interviews from the cell phone sample.

All CARAVAN® interviews are conducted using ORC International's (ORC) computer assisted telephone interviewing (CATI) system.

As required by the Code of Standards of the Council of American Survey Research Organizations (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.

Sampling

The CARAVAN® landline-cell combined sample is a dual frame sampling design. This means that the sample is drawn from two independent non-overlapping sample frames—one for landlines and one for cell phones.

Land Line Sample

ORC's Random Digit Dial (RDD) telephone sample is generated using a list-assisted methodology. That is, the updated white page listings that are used to identify telephone number banks (the first 8 digits of the phone number) with a listed phone number in them. The standard that we use is 2+, meaning that a bank needs to have 2 or more listed households to be considered working. We use the Genesys Sampling in-house system to generate list-assisted Random Digit Dialing sample.

The standard GENESYS RDD methodology produces a strict single stage, EPSEM (Equal Opportunity of Selection Method) sample of residential telephone numbers. In other words, a GENESYS RDD sample ensures an equal and known probability of selection for every residential telephone number in the sample frame.

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, switch locations for the selected thousand series blocks are determined. This provides the information needed to map each cellular thousand series block to a county.

The cell phone sample was generated from cell phone only 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.

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 as well as 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.

For CARAVAN[®] (landline), the post survey weighting adjustments leverage population-based estimates as reported by the 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 an iterative 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. For CARAVAN[®], the ratio adjustments calibrate the survey data to the population for age, sex, race/Hispanic origin, Census region, and education.

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 three subgroups:

- Landline mostly*: those who receive most calls on a landline,
- True dual*: those who receive calls on both regularly, and
- Cell mostly*: those who receive most calls on a cell phone.

The National Health Interview Survey (NHIS) provides estimates of these user group populations. The NHIS is an in-person survey conducted by the National Center for Health Statistics (NCHS). We weight-adjust the landline sample and the cell sample to their respective population estimates from the NHIS.

Since the dual user groups are represented by the landline sample and cell phone sample, we combine these groups based on a weighted average. The weighted average is based on the effective sample sizes for each group.

Finally, the combined sample is weighted to represent the US population using data from the Current Population Survey (CPS) on age, gender, race, region, and education as noted above.

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*.