

Technical Information

Country	Year	Sample Size	Weighted/Unweighted	Fieldwork dates
---------	------	-------------	---------------------	-----------------

protect against outsourcing work to untrained interviewers, and assessing the timing of the interviews.

For the 2016/17 round, LAPOP introduced a new quality control system to ensure the highest practicable fieldwork quality: Fieldwork Algorithm for LAPOP's Control over Survey Operations and Norms (FALCON©). To build capacity, staff in the offices at each local firm were trained in FALCON procedures and collaborated with LAPOP over their implementation. The system was refined as we progressed through the 2016/17 round, and in its final form is composed of the following elements:

1) Geo-fence Module: LAPOP's Geo-fence Module ensures that interviews are collected in the assigned work area through the creation of a series of circular boundaries placed around selected census segments (when census bureaus make the shape file available) or municipalities (which are the geographic areas that LAPOP typically uses as the Primary Sampling Unit). Using ArcGIS, LAPOP calculates GPS coordinates for each Primary Sampling Unit's unique centroid using the shapefiles available for each country¹ and estimates its Minimum Bounding Geometry (MBG), that is, the smallest possible circle drawn around the shape file for each selected municipality or census segment.² The Geo-fencing Module then flags all interviews conducted beyond the fences, by an automatic tool programmed into the data collection software (see the Distance Audit Module below). Interviewers are notified by their devices if the GPS coordinates captured lie outside the geo-fence (meaning that the interview may be in the wrong location). At this point, interviewers are instructed to move to the correct location and/or notify their supervisors of the problem and request corrective information. Since, however, errors can occur in the programming or collecting of GPS data and because interviewers sometimes need to proceed out of the geo-fence because of conditions on the ground, the system has a built-in flexibility that allows the interviewer to proceed, but only *after* formally indicating to the FALCON system on their Android device, that they are knowingly in the wrong geo-referenced location. Once an interview has been completed, the system automatically uploads it to the cloud, and it becomes instantly visible by the fieldwork supervisors in the given country, as well as by LAPOP Central. Corrective instructions can then be communicated to each interviewer and/or field supervisor. This entire process is automated and made highly efficient by the Distance Audit Module, described below.

2) Distance Audit Module: LAPOP's Distance Audit Module (DAM) provides assessments of interviewers' distance from the bounds of the geo-fence. Once each interview is uploaded, the team monitoring fieldwork is able to determine whether an interview was carried out in the

in kilometers (and fractions of a kilometer) between the interviewer's location at the moment of the interview and the closest point of the circumference around the census segment or municipality (i.e., the limit of the geo-fence). If the flag shows a minor variation, this might be a result of inaccurate GPS data. In other circumstances, DAM may uncover unintentional or intentional errors on the part of the field staff that would lead the supervisors or LAPOP auditors to cancel the errant interview.

3) Location Consistency Check: FALCON's Location Consistency Check (LCC) assures that interviewers are in the correct (i.e., designated) location *before* each interview takes place. If the location of interviewers is not the one assigned by fieldwork supervisors, the software immediately informs the interviewer of the problem so that it can be corrected. The interviewer is not allowed to proceed if the sample segment assigned by the home office is located in a municipality (and district) different from the one indicated by the interviewer. The LCC thus helps ensure that interviewers collect data from the location selected in the sample and not from another community with an identical or similar name.

4) Multi-Tiered Auditing: In addition to the checks highlighted above, the SurveyToGo software is programmed to collect additional information that allows monitoring both the quality of interviews and the identity of interviewers. With respect to the quality of interviews, LAPOP silently records a subset of questions over the course of each interview in order to assess if the survey questions are being read appropriately.³ Additionally, LAPOP times the net duration of the interviews to determine if they are being carried out in a reasonable timeframe. Concerning the identity of interviewers, the software silently captures photographs of the interviewer, and collects their signatures with the purpose of providing evidence that the person gathering data is the one LAPOP trained and certified as interviewer. Once this information is in the system, fieldwork teams listen to the recordings, check the photographs, review the net durations, and verify the signatures from 100% of interviews to assure that enumerators adhered to best survey practices. If interviews comply with LAPOP quality standards, they are initially approved; otherwise, they are canceled. A second level quality control team audits a random subset of initially approved interviews to assure the quality of interviews and the quality of field teams' checks. Following this protocol allows LAPOP projects to provide quality feedback to interviewers and field supervisors in real time, correcting errors, coaching interviewers to read more clearly or slowly, canceling and replacing low-quality interviews, and giving appropriate recognition to high-quality work.

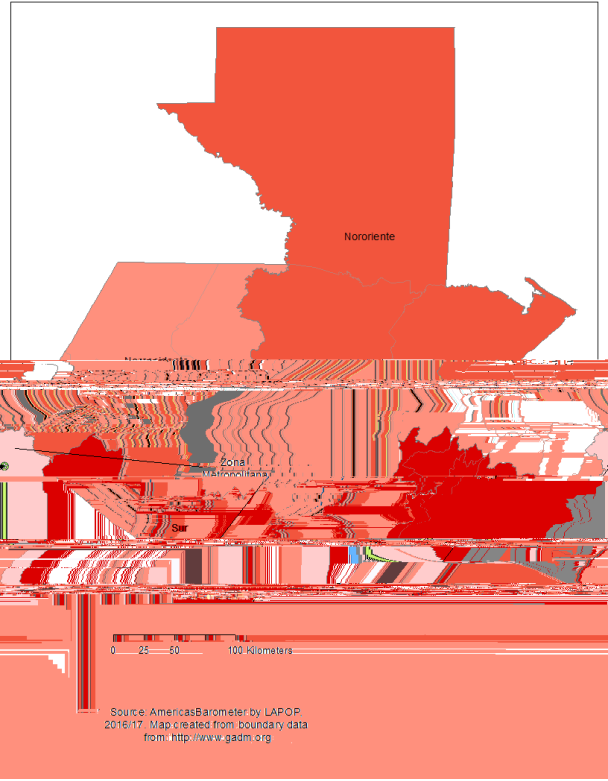
As per the sample design, the 2016/17 round of the AmericasBarometer continues to use the sample strategy introduced for the first time in the 2012 round of the surveys and that was also employed in 2014. This sample design continues to use, in almost all cases, the same stratification employed since 2004, making adjustments where necessary when census information is updated. The samples are all representative at the stratum level. The new design, however, stabilized the PSU and cluster sizes, with the selection of each PSU based on PPS (Probability Proportional to

³ Interviewers are informed in training that their voices would be recorded, though not told which sections would be captured. Respondents, in the information instructions read to them before the interview began, are likewise told that portions of the interview would be recorded for quality control purposes.

Size). Within PSUs, clusters are also standardized (typically 6 interviews) to minimize intra-class correlation while taking advantage of economies of fieldwork that simple random selection of interviews within the entire PSU would not make possible.

The tradeoff continues to make the sample design very efficient with very low intra-class

The sample consists of 62 primary sampling units and 251 final sampling units including all departments in Guatemala. A total of 758 respondents were surveyed in urban areas and 788 in rural areas. The estimated margin of error for the survey is ± 2.5 . The margin of sampling errors is not adjusted for weights. Table 1 shows the sample size in each of the nine regions (strata) and by municipality size.



Strata	Unweighted Sample Size
Metropolitan Area	398
Northeastern	445
Northwestern	268
South	435
Total	1,546
Size of Municipality	
More than 75,000 inhabitants	267
Between 25,000 and 75,000 inhabitants	488
Less than 25,000 inhabitants	791
Total	1,546

LAPOP uses “frequency matching,” a technique that permits one to obtain a sample with similar distribution of age and gender to that of the national census or electoral registration lists. Frequency matching avoids the extremely costly effort involved in making multiple callbacks to each missed unit within each PSU in an effort to obtain a balanced sample. In national, face-to-face interviewing, multiple callbacks are often impractical from a cost standpoint. Our experience reveals that even three callbacks leave the sample with a notable gender imbalance (more women than men, since women are more likely to be at home than men). Rather than having to include post-hoc weights to adjust for this sample error, which can be large, we resolve the problem in the field via using a distribution of interviews among gender and ages that reflects the structure of the population.⁶

A single respondent was selected in each household, following the frequency matching distribution programmed into the sample design, by gender and age as mentioned above. Respondents are limited to household members who reside permanently in that household (thus excluding visiting relatives), who fit the age and residency requirements (limited to adult citizens and permanent residents). If two or more people of the same sex and age group were present in the household at the moment of the visit of our interviewer, the questionnaire was applied to the person who most recently celebrated a birthday (i.e., the “last birthday” system) in order to avoid selection bias.

Geo-fences were programmed at the municipal level in Guatemala and compliance reviewed on a daily basis to assure that interviews took place in the correct location.

The dataset contains a variable called “wt” which is the “country weight” variable. Since in the case of Guatemala the sample is self-weighted, the value of each case = 1. When using this dataset for cross-country comparisons, in order to give each country in the study an identical weight in the pooled sample, LAPOP reweights each country data set in the merged files so that each country has an N of 1,500. The weight variable for cross-country comparisons is called “weight1500”. In SPSS, this is done via the “weight” command. Weights are already activated in SPSS datasets. In Stata, one should use the svyset command to weight the data and declare the sampling information to correctly compute standard errors that take into account the design effects. The command for single country, single year studies is: **svyset upm [pw=wt], strata(estratopri)**. For cross-country and/or cross-time studies, the command is: **svyset upm [pw=weight1500], strata(strata)**. These declarations have been made in Stata datasets. However, you must use the svy prefix with estimation commands to compute the weighted statistics and correct standard errors (see help svy_estimation within Stata for more information).

For additional information contact Georgina Pizzolitto at georgina.pizzolitto@vanderbilt.edu.

⁶ An alternative strategy is to post hoc weight such samples in order to force the sample to correspond to the census distributions. However, if the fieldwork produces a substantial deviation from those distributions, the result could be placing excessive confidence on a very small number of respondents for some population group (e.g., older males). The resulting widening of confidence intervals for these weighted small sample group could limit inferences drawn from such weighted samples.