VANDERBILT

LAPOP

AmericasBarometer
Barómetro de las Américas

AmericasBarometer, 2016/17

Technical Information

Country	Year	Sample Size	Weighted/Unweighted	Fieldwork dates
Brazil	2017	1,532	Weighted	April 5th – May 11th

LAPOP AmericasBarometer 2016/17 round of surveys

The Latin American Public Opinion Project (LAPOP) is a pioneer in survey research methods. In the 2016/17 round of the AmericasBarometer, LAPOP has continued this tradition of innovation, with heretofore unprecedented improvements in monitoring interview quality on a daily basis during the course of fieldwork. This was done by making significant advances in the use of handheld and expansion of electronic devices for data collection, coupled with a wide variety of new quality control techniques utilizing LAPOP's FALCON© system (see details below).

In the 2016/17 round of the AmericasBarometer, handheld devices for data collection were used in 100% of the countries surveyed, for all interviews. The sole exception is Haiti, where approximately 50% of interviews were conducted using paper questionnaires, a choice dictated by scarcity in data signals, internet connections and power to recharge devices. As in prior rounds of the AmericasBarometer, the U.S. and Canada studies were conducted online while all other interviews were conducted face-to-face. In the 2016/17 round, we predominantly used the SurveyToGo© (STG) software (as an exception, we used Adgys in the 6 OECS countries), running on Android tablets and phones, to conduct field interviews. LAPOP has found this software to be very reliable and flexible. Importantly, the adaptable platform and accessible programming language has allowed LAPOP to program in numerous customized modules that enable our quality control protocols.

The use of electronic devices for interviews helps us improve efficiency in data collection in several ways. First, it eliminates data entry errors that occur when handwritten responses are transferred to digital formats by coders and data entry clerks. Second, it supports user-friendly switching among multiple languages, especially important in countries like Paraguay, in which large proportions of respondents code-switch between Spanish and Guaraní. Third, it provides quality control teams the ability to audit and track the progress of fieldwork on a daily basis. The LAPOP auditing and tracking process includes verifying that interviews are being carried out in the pre-selected sampling locations, ensuring the correct and precise reading of the full wording of questions and response choices, checking the identity of interviewers for each survey to

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 country[1] 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 correct place, and if not, whether the violation was major or minor. The DAM returns the distance

[1] Shapefiles used for each country were obtained from Census Bureau and/or Government websites or, when those were not available, from public access geospatial websites.

2 An exact tracing of the boundaries of each shapefile is impossible given enormous number of very precise coordinates that would have to be included in the sample data base. The MGB provides the best approximation of the land area within each shape file. In practice, this means that while 100% of all segment or municipal area land area is always included within the geo-fences LAPOP draws, the circles will include areas beyond the boundaries of the segment/municipality that will register as "within-boundary" when in fact they are beyond the official boundaries of the shape files.

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 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 correlations. With the cluster and PSU sample size uniformity, the LAPOP samples are now also representative within each selected municipality, to enable the use of the municipality as a unit of analysis for multilevel statistical analysis. However, with the small sample sizes at the PSU level that our design produces, confidence intervals at the level of each PSU are, by definition, wide. Users of the data should note that while the stratification incorporates all major regions of the country (exceptions include islands, such as the Galapagos in Ecuador or San Andrés in Colombia, but they do include the Bay Islands of Honduras), and therefore can be reliably used to analyze differences among strata, the PSUs selected normally represent only a small fraction of the total PSUs in the country (typically 50-65). Details of the sample design revisions are found in the description of the 2012 AmericasBarometer surveys.

The remaining pages of this technical note describe the sample design of the 2016/17 AmericasBarometer survey in Brazil.

Brazil 2016/17 AmericasBarometer Survey

This survey was carried out between April 5th and May 11th of 2017, as part of LAPOP's AmericasBarometer 2016/17 wave of surveys. It is a follow up to the national surveys of 2006 2008, 2010, 2012 and 2014 carried out by the Latin America Public Opinion Project (LAPOP). The 2016 survey fieldwork was carried out by IPSOS Brazil on behalf of LAPOP. Key funding came from Vanderbilt University and the Open Society Foundation.

The project used a national probability sample design of voting-age adults, with a total N of 1,532 people involving face-to-face interviews conducted in Portuguese. The survey used a complex sample design, taking into account stratification and clustering.

The sample was developed by LAPOP, using a multi-stage probability design and was stratified by the five major regions of the country: North, Northeastern, Mid-west, Southeastern and South. The map below shows the geographic division of these strata (see Figure 1). Each stratum was further sub-stratified by size of municipality[4] and by urban and rural areas within municipalities. Respondents were selected in clusters of 6 in urban and rural areas. Reported statistics or statistical analyses should be adjusted for the design effect due to the complex design of the sample.[5]

[4] The new sample design included five different strata of municipalities classified according to their size. Municipalities were grouped in sizes as follow: (1) municipalities with less than 25,000 inhabitants, (2) medium-sized municipalities with between 25,000 and 100,000 inhabitants (3) large municipalities with more than 100,000 inhabitants.

[5] For more information visit http://www.vanderbilt.edu/lapop/survey-designs.php

The sample consists of 125 primary sampling units and 250 final sampling units including all states in Brazil. A total of 1,322 respondents were surveyed in urban areas and 210 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 five regions (strata) and by municipality size.

Figure 1: Sample stratification in Brazil

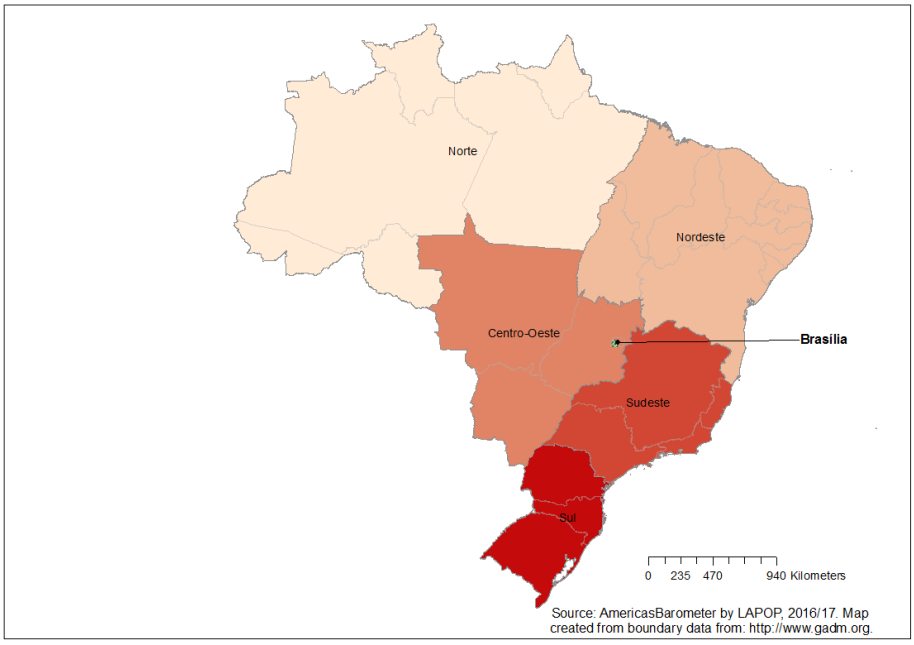

Table 1: Sample sizes by Strata and Municipality Size in the 2017 AmericasBarometer Survey in Brazil

Strata	Unweighted Sample Size
North	219
Northeastern	346
Mid-west	217
Southeastern	491
South	259
Total	**1,532**
Size of Municipality	
More than 500,000 inhabitants	531
Between 50,000 and 500,000 inhabitants	526
Less than 25,000	475
Total	**1,532**

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 Brazil and compliance reviewed on a daily basis to assure that interviews took place in the correct location.

Weighting of the Brazil datasets

The dataset contains a variable called "wt" which is the "country weight" variable. Since in the case of Brazil the sample is weighted, the variable "wt" must be used in the estimations. 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.