

2015 ARGENTINE PANEL ELECTION STUDY

SAMPLE DESIGN

APES conducted two national, face-to-face waves of interviews, the first between June 24, 2015 and August 7, 2015, and the second between November 21, 2015 and December 30, 2015. The first wave was based on a nationally representative sample of Argentine voters living in cities of 10,000 inhabitants and more, while the second wave consisted of a panel sample of those wave 1 respondents that accepted to participate again in the second wave, plus a refresh sample.

Wave 1 Sampling

The first wave of the APES relied on a national household sample of 1,149 Argentine citizens aged 18 years and over. The general design was a stratified multistage cluster sample, using probabilistic selection at some stages and non-probabilistic selection at others. Given the non-probabilistic design elements and the non-probabilistic nature of all human samples (because of non-random participation patterns), estimating sampling error is not appropriate. As a point of reference, a simple random sample of 1,149 cases would result in a standard error of approximately $\pm 1.5\%$ for a sample proportion of 50%. A 95% confidence interval for such a proportion, therefore, would be 47-53%.

Table 1 summarizes the details of the multistage sample design (the stratification rationale and procedure are explained below).

Table 1. APES multistage sampling

Stage	Sampling unit	Number of units	Cases within unit			Selection procedure
			Mean	Min	Max	
1	Cities/towns	18	63.83	27	349	Purposive and convenience sampling within strata
2	Census fractions					PPS sampling
3	Census radiuses					PPS sampling
4	Blocks/PSU	160	7.18	4	9	Simple random sampling
5	Households	1,149	1	1	1	Systematic random sampling
6	Individuals	1,149	1	1	1	Gender and age quotas

Note: PPS stands for “Probability Proportional to Size”

Stages 2 through 6 were conducted according to standard household sampling procedures, using different types of probabilistic selection procedures when the population is composed of geographic areas or households, and turning to a quota (non-probability) procedure at the last stage, where the population is individuals. The high cost of probability sampling at this stage made it advisable to turn to standard quota procedures, which are easier to implement, less expensive, and less disruptive of the relationship between the interviewers and the members of the selected households. A combination of gender and age-range quotas were used at the PSU level, so that interviewers in a given block had to look for specific types of interviewers in order to complete the assigned gender and age profiles.

The selection of cities in stage 1 followed a different strategy. Cost consideration required us to select only 18 localities at this stage, a relatively small sample. As randomness requires large numbers to produce representative samples, the selection of cities was done with alternative methods. We first divided the country's 23 provinces (plus the federal district of Buenos Aires) into four electoral strata. That is, instead of stratifying the sample by the usual criteria in public opinion surveys (e.g., region or city size), we stratified according to our main variable of interest, electoral behavior. This was especially appropriate in a context in which politics and elections have become denationalized and regional.

The APES team constructed four strata based on the 2011 presidential electoral results, using as the stratifying variable the percentage of the valid votes obtained by the incumbent Cristina Fernández de Kirchner. Provinces were classified into four strata and cities/towns selected within each stratum. Table 2 presents the details.

Table 2. *APES stratification by prior vote*

CFK 2011 vote	Substratum	Weight	Province	City/town	Effective N			
<45%	Federal Capital	10.27	Capital Federal	Capital Federal	117			
	Córdoba San Luis Santa Fe	16.15	Córdoba	Córdoba Capital	50			
			Córdoba	Alta Gracia	46			
			Santa Fe	Rosario	46			
			Santa Fe	Santa Fe	39			
45–63%	Greater Buenos Aires	29.98	Buenos Aires	Greater Buenos Aires	349			
	Buenos Aires (int) Chubut Entre Ríos La Pampa La Rioja Mendoza Neuquén	23.97	Mendoza	Gran Mendoza	48			
			Buenos Aires	Mar del Plata	46			
			Entre Ríos	Paraná	38			
			La Rioja	La Rioja	49			
			Neuquén	Neuquén	46			
			Buenos Aires	Ayacucho	47			
			63–72%	Catamarca Corrientes Chaco Jujuy Misiones Río Negro Salta San Juan Tucumán Tierra del Fuego	16.89	Tucumán	Gran Tucumán	47
Chaco	Resistencia	39						
Corrientes	Corrientes	29						
Jujuy	S. S. de Jujuy	48						
Río Negro	Cipolletti	27						
>72%	Formosa Santa Cruz Santiago del Estero	2.75				Santiago del Estero	Termas de Río Hondo	38
		100.00			1,149			

The selection of cities within each stratum was based on a combination of purposeful and random sampling. All the cities larger than 10,000 inhabitants in each electoral stratum were listed. From this sampling frame the APES team randomly selected cities, but ensuring

diversity in terms of size (such that the sample includes the range of cities from the metropolis of Buenos Aires, to large cities such as Córdoba or Mendoza, to medium-sized cities such as Corrientes and La Rioja, to small cities such as Cipolletti and Ayacucho) and in terms of the province in which they are located (giving more weight to larger provinces such that some are represented by two cities (e.g., Buenos Aires, Córdoba) and some by zero cities (e.g., Catamarca, Tierra del Fuego). Finally, cost considerations meant avoiding remote cities (e.g., in Southern Patagonia).

The metropolitan area of the city of Buenos Aires (where 32% of Argentines live), was given special treatment. Since it stretches over two different subnational districts (the federal capital and the province of Buenos Aires), it was divided into the City of Buenos Aires (or federal district) and Greater Buenos Aires. These two areas were defined as substrata within their strata and included in the sample with a number of cases proportional to their weight in the population (that is, in the universe of Argentines living in cities of 10,000 inhabitants or larger). This means that about 40% of the country's adult population is fully represented by the inclusion in the sample of the City of Buenos Aires and Greater Buenos Aires, while the remaining 60% is represented by the other 16 cities.

Wave 2 Sampling

The panel design implied that we attempt to re-interview all respondents from wave 1. Our success rate was of 68% (780 out of 1,149 original respondents). To compensate for sample attrition, we drew a refresh sample of 626 respondents, selected according to the same procedures described above for wave 1. The wave 2 sample therefore has a sample size of 1,406.