Elections polling – Sample and methodology

Traditionally telephone surveys have been conducted by landline. In Afghanistan however, as is the case in many developing countries where the large scale uptake of communication technologies is relatively recent, the proliferation of mobile phones is far greater than that of landlines. The telecommunication network in Afghanistan has expanded exponentially in the past decade (currently to an estimated 17 million mobile phone subscribers[1] thus providing the opportunity to engage with individuals and households not only in the urban centers, but also in the most remote areas of the country

All calls were made from a central location in Kabul, by the same group of surveyors during the entire length of the study. All calls were dialed manually.

There are no up to date registries of mobile phones in Afghanistan; therefore, the sampling method for telephone interviewing is reliant on utilizing an algorithm to generate a random list of mobile telephone numbers and working through those numbers individually. Such a method is time consuming; however it is the only means to ensure a truly impartial selection of respondents.

The telephone sample is representative of the Afghan population possessing and/or who can be reached by phone: since it is common to have one mobile phone for the entire household, the initial respondent to the telephone call (usually a man) was not necessarily the one interviewed, often surveyors requested the phone be passed to female members of the household who were then interviewed. .

As a result of pervasive electoral fraud in Afghanistan the number of registered voters in the elections or the number of voting cardholders does not reflect the actual amount of 'likely voters' in the country. The sample in this poll was thus selected from all citizens of voting age, which also served to give voice to those citizens who for security or accessibility or other reasons may not be able to access polling stations.

The sampling methodology minimized the risk of non-responses (if the phone number is not working or out of coverage; if the respondent is not picking up; or if the respondent does not want to answer) in the following manner:
- Afghans (both men and women) tend to respond more positively to surveys than in Western countries, notably because there is less "phone survey fatigue" in Afghanistan than in other countries.
- Whenever the line cuts or the quality of the call were inadequate, surveyors recall the respondent in order to obtain their response
- In cases where the respondent does not pick up the phone, surveyors attempt to recall them at another time of the day

---

[1] Afghanistan's Ministry of Communication & Information Technology

- The question is succinct and formulated in the simplest and most accessible terminology to minimize risks of any misunderstandings
- Surveyors call from phones with no special access code (such as toll-free numbers), and that do not generate automatic number identification; therefore increasing the probability that the interviewee picks up the phone.

ATR consulting randomly called 3,222 people across the country. The margin of error for the overall sample (a measure of the approximation in the results inherent when using a representative sample) is less than 1%. The confidence level for the poll is 95%, meaning that there is a 95% probability that the results are consistent with the results that would be obtained should the survey be conducted with the entire target population.. The margin of error is greater for subgroups (women). However, the sample is of an adequate size to allow for statistically significant conclusions to be drawn from the subgroup analysis.

The following question was asked of each respondent: *"Which candidate in the presidential elections are you going to vote for?"* The open formulation of the question negates any potential to lead the respondent, ensuring maximum comparability with the results of the previous polls.

People could either select one of the 11 candidates who remain in contention or respond *'I do not know'* or *'None of them*.

The survey took place from 19 March to 26 March 2014, ten days before the tenure of elections. The timing of data collection is crucial to interpreting results.

A. Training surveyors

All surveyors have prior experience with ATR's methods of data collection and uncompromising standards of data quality. In addition, before data collection commenced all surveyors received a day long refresher course, providing an overview of the project, challenges specific to election polling (such as the sensitivity of the topic, normative responses, concerns over anonymity etc.) and means to overcome such. The majority of surveyors also had prior experience with election polls specifically and thus were familiar with the issues involved. For security reasons, ATR commits to not disclose the names of the female surveyors working on this project.

During the training, ATR presented to surveyors a strict protocol to be followed in all interviews. This included the manner in which they introduced themselves and the exact wording to be used in formulating the question. Surveyors were required to name ATR consulting as the company conducting the survey and the overall purpose of the research when requested by respondents. Surveyors did not use their real name on the phone in order to protect their identities.

All surveyors possessed the exact wording of the question in both Dari and Pashto in front of them at all times so as to minimize any potential variance. In addition all surveyors were intensively trained on the importance of impartiality. A full time team leader was assigned to monitor calls, guide the surveyors

and provide feedback where needed. During data collection, the team leader reported to the project manager (see Monitoring section below).

During the training surveyors recorded socio-economic data of all respondents, including whether the respondent originated from an urban or rural area. As regional quotas were applied, all surveyors referred directly to a map of the country and a list of provinces belonging to each region.

The question was open, in that the surveyors were directed not to read the list of candidates. However in cases where the respondent was not aware of the names of all candidates, surveyors read the entire list in alphabetical order.

As elections draw closer, the sensitivity of the topic is expected to increase, as is respondent's unwillingness to discuss voting preferences on the phone. In such cases surveyors are trained not be coercive, rather to reiterate the purpose of the survey and the conditions of anonymity. If the respondent still does not wish to participate surveyors would thank them for their time and terminate the call.

During training, surveyors were presented with various scenarios based on ATR's polling experience. For example:

Case 1: If the respondent answers:
- *"I do not know who the candidates are"*
The surveyor should then read the entire list of candidates, in alphabetical order, based on the Dari alphabet.

Case 2: If the respondent answers:
- *"This is none of your business", "I do not want to respond",* or *"I have not decided yet"*
Surveyors have to record it in the category *'I do not know'*.

Case 3: If the respondent answers:
- *"Karzai"*
If people respond *'Hamid Karzai'* or *"the current president",* the surveyor should inform the respondent that he is not a candidate and ask whether the respondent would like the list of candidates to be read to him / her. (Then refer to case 1)

Case 4: If the respondent answers:
- *"The one Karzai supports"*
In that case, the surveyor must ask the respondent to name a candidate in the list which they think Karzai will support.

Case 5: If the respondent names a previous candidate who is not on the final list of candidates anymore or (for instance, *"Abdul Rahim Wardak"*) or names a politician who is not candidate (for instance, *"Yunus*

Qanuni"), surveyors inform the respondent that he is not a candidate and ask whether the respondent would like the list of candidates to be read to him / her. (Then refer to case 1)

Case 6: If the respondent answers:
- *"All are bad", "I don't want to vote for any of them", "I will not vote",* or anything similar
Surveyors record the response as *"none of them".*

Case 7: If the respondent misunderstands the question and responds *"I do not vote in the Parliamentary elections"*, the surveyors repeat clearly that the survey concerns presidential elections.

Case 8: If the respondent is under eighteen years old (the legal age to vote), surveyors should ask the respondent to pass on the phone to a family member aged 18 or above to answer the survey.

Case 9: If a respondent answers:
- *"I do not have a voting card", "I am not registered to vote"* or *"I do not know how to vote",* surveyors are told to select *"None of them".*

Case 10: If the respondent is defensive and asks the surveyor to tell her first who she is going to vote for, surveyors do not give their personal opinion and should explain why for reasons of impartiality they are not permitted to do so.

During the training, each surveyor is required to make several dummy calls under supervision during the practical sessions to ensure that all protocols are respected. Surveyors receive feedback on each call in a group setting and take note of each other's mistakes.

The monitoring system is explained to the surveyors so they are fully aware of the checks and balances to detect any false data entry. Surveyors are also all made aware that that any attempt to enter false data will result in immediate dismissal.

B. Selecting respondents

ATR utilized an algorithm to generate more than 20 million mobile phone numbers at random. A list of 10,000 phone numbers randomly selected from the full lists distributed to each of the 16 surveyors. Surveyors then worked through the list to determine which numbers were attributed, interviewing the respondents who were willing to give their feedback.

To ensure the highest possible degree of access to female respondents all surveyors were women. Techniques to increase access to female respondents were also used such as asking a male respondent if there were any female members of their family in the vicinity who would be willing to respond to the survey. This requires surveyors to be extremely polite in asking to speak with a female member of the respondent's family. In this way the survey is representative not only of individuals with phones, but also of those households with one phone for all family members.

Surveyors registered any respondents they could reach over the phone, up to the time they reached a specific target within the list of quotas (see below). Once a quota is reached, surveyors cannot register a respondent who falls into this category anymore.

C. Designing the sample and fixing quota

In order to ensure that the sample for each of the five regions (see below table 1 for a presentation of the regions) is representative, a total of 384 respondents were interviewed in each of these regions.

Table 1: Regions

Region	Provinces
Central	Kabul, Wardak, Logar, Parwan, Kapisa, Panjsher, Bamyan, Daikundi
North	Badakhshan, Takhar, Kunduz, Baghlan, Samangan, Sar-i-Pul, Jawzjan, Balkh, Faryab
South	Helmand, Kandahar, Uruzgan, Zabul, Ghazni
East	Paktya, Paktika, Khost, Laghman, Nangarhar, Kunar, Nuristan
West	Herat, Badghis, Farah, Ghor, Nimroz

Table 2: Sample for the Central region

Central region					
Kabul city		Rural		Urban	
Male	Female	Male	Female	Male	Female
352	201	156	57	63	13

Table 3: Sample for the North region

North region			
Rural		Urban	
Male	Female	Male	Female
165	77	297	118

Table 4: Sample for the East region

West			
Rural		Urban	
Male	Female	Male	Female
164	103	177	126

Table 5: Sample for the South region

South			
Rural		Urban	
Male	Female	Male	Female
207	41	253	62

Table 6: Sample for the East region

East			
Rural		Urban	
Male	Female	Male	Female
227	91	198	74

The confidence level is 95%. The confidence interval for each category is presented in the below table:

Table 7: Sample size and confidence intervals per gender, area of residence and regions

Demographic	Sample Size	Confidence Interval	Margin of Error (+/- %)
Men	2258	2.06	1.03
Women	962	3.16	1.58
Rural	1288	2.73	1.37
Urban	1933	2.23	1.12
Kabul City	553	4.17	2.09
Central Region (including Kabul districts)	289	5.76	2.88
North Region	657	3.82	1.91
West Region	570	4.1	2.05
South Region	563	4.13	2.07
East Region	590	4.03	2.02

D. Monitoring

A team leader monitored the surveyors during the entire data collection period. The team leader listened to conversations in real time while data was being entered, selected random entries from each surveyor and called back respondents to verify they have been interviewed and oversaw the entry of all respondents into the database. The team leader was also monitoring the attitude and phone manner of all surveyors and ensured that questions were delivered consistently. The team leader reported twice a day to the project manager (before lunch and at the end of the day) on progress towards quotas and attitude of surveyors.

In addition all surveys were recorded and two separate monitors (one Afghan and one international who understands Dari and Pashtu) each showed up regularly to monitor the surveyors and the team leader. They listened to approximately 10% of all surveys each (covering a total of 20% of surveys) selected at random to ensure none of the surveyors were influencing the responses and data was of an acceptable quality.

E. Data cleaning

In addition to the monitoring system implemented during the polling, ATR performed data cleaning in order to comply with standard quality procedures. Surveys were discarded in the following situations:
- Two same phone numbers were entered by mistake in the database
- One of the questions was not completed and the cell left blank
- The age recorded is below 18 years old or includes a typo ('235' for instance)

The use of a database with limited pre-listed options reduced to zero the chances for surveyors to enter responses that differed in spelling or in meaning.

F. Methodology of the analysis

Poll results are weighted in order to account for unequal accessibility of different respondents. Firstly the results were weighted within each province to ensure equal representation of male, female, urban and rural respondents. Nationwide results were then reached by weighting results by province according to the national population distribution. ATR uses population data from the latest report from the Central Statistics Office.

G. Principles of disclosure

ATR consulting abides by the following principles of disclosure:
- ✓ Sponsorship of the survey
- ✓ Fieldwork provider
- ✓ Dates of interviewing
- ✓ Sampling method employed

- ✓ Sample population
- ✓ Size of the sample that serves as the primary basis of the survey report
- ✓ Margin of sampling error
- ✓ Survey mode
- ✓ Complete wording and ordering of questions
- ✓ Percentage results of all questions reported