DI 2008 Sampling Design

Overview

The target population of the 2008 DI encompasses all households in Armenia, Azerbaijan, and Georgia, except those in the conflict areas (Nagorno-Karabakh in Azerbaijan, Abkhazia and South Ossetia in Georgia). The survey followed a stratified two stage sampling design.

Stratification

To ensure representativeness and to preserve compatibility with previous versions of the survey, the 2008 DI survey used 9 strata. Each country was divided into 4 geographical quadrants and the capital. Each of the 4 non-capital quadrants was divided into urban and rural strata. The number of PSUs in each stratum were selected proportional to the population of each stratum, according to census data in Azerbaijan and Georgia and electricity records in Armenia.

Table 1: Proportion of the Population in Each Stratum

Stratum	Armenia	Azerbaijan	Georgia
Capital	0.35	0.35	0.31
North East Urban	0.09	0.1	0.06
North East Rural	0.11	0.04	0.11
North West Urban	0.13	0.12	0.13
North West Rural	0.1	0.086	0.13
South East Urban	0.03	0.15	0.07
South East Rural	0.03	0.13	0.07
South West Urban	0.05	0.04	0.06
South West Rural	0.1	0.01	0.05

Primary Stage

Primary Sampling Units

The primary sampling units are voting precincts in the cases of Azerbaijan and Georgia, and electricity grid groups in Armenia. In Azerbaijan and Georgia, the precincts were selected with probability proportional to the number of registered voters assigned. In Armenia, PSUs were all basically equal in size, so simple stratified random sampling was used.

Second Stage

Secondary Sampling Units

Armenia

Because of the expected accuracy of the electricity data in Armenia, households were sampled directly from electricity grid records.

Azerbaijan

In Azerbaijan, due to costs constraints, households were sampled in each precinct using "random walk" method.

Georgia

In Georgia, after PSUs were selected, enumerators block listed selected precincts using hand-drawn maps when necessary. After block listing, enumerators created a sample frame of households in each PSU, comprising all households living in the PSU. From this sample frame, households were randomly sampled.

Sample Size

To calculate sample sizes for each country, CRRC balanced the need for large samples for acceptable precision with the considerable costs involved in conducting face-to-face interviews. The original goal for the survey was to have sample sizes sufficient for estimates (for binary variables) with margin of errors of ±5% with 95% confidence intervals in each "macro-strata", defined as rural, urban, and capital households. For Armenia and Georgia, sufficient number of PSUs and households were sampled to achieve this goal. In Azerbaijan, however, the sampling design only allowed the desired level of precision at the national level, not for each "macro-strata".

To calculate the required sample sizes, we require a measure of how cluster sampling affects the precision of the sampling design (the design effect), and an estimate of the nonresponse rate. These measures are discussed below.

Design Effects

Because of the two-stage design, sample size calculations must take into account the loss of precision resulting from cluster sampling. To adjust the sample sizes for clustering, ideally a "design effect" would be estimated using previous survey results and the sample could be adjusted by factor equal to the design effect. For Armenia, given that the 2007 survey used the same basic design as the 2008 survey, this adjustment is straightforward. For Azerbaijan and Georgia, however, the primary sampling unit was changed from census tract to electoral precinct, making design effects not directly comparable across the two surveys. As a result, some arbitrary decisions are unavoidable in the following sample size calculations.

Despite this important limitation, design effects using the 2007 DI data were calculated for all 3 macro-strata in all 3 countries, to at least give a general sense of how much sample sizes were to be adjusted.[1] Note that these design effects are for clusters with 50 respondents sampled in each PSU. The results of these calculations are presented in Table 2.

[1] An average of the design effects for the following three variables were used: proportion of population over the age of 60, proportion of population at each education level, and self-assessed health.

Table 2: Design Effects - 2007 DI Survey

Macro-strata	Armenia	Azerbaijan	Georgia
Capital	2.19	2.42	3.99
Rural	2.91	3.96	2.18
Urban	1.44	3.45	2.64

The design effects calculated for Table 2 assume 50 attempted respondents per PSU. Initial sample size calculations from the 2007 DI survey assuming 50 respondents per PSU and using these design effect assumptions resulted in sample sizes that were very large and too costly. Calculations using the formula for the variance of cluster samples showed, however, that very little precision would be lost by decreasing the number of respondents per PSU. Moreover, cost savings from fewer respondents per cluster could be used to increase the number of PSUs sampled, and the desired level of precision could be maintained.

Nonresponse Rates

Nonresponse rates were calculated using the 2007 DI survey for the three macro-strata in each country. They are presented below in Table 3.

Table 3: Non-Response Rates - 2007 DI Survey

Macro-strata	Armenia	Azerbaijan	Georgia
Capital	0.7	0.73	0.72
Rural	0.83	0.68	0.86
Urban	0.79	0.76	0.72

Final Sample Size Calculations

Simulations[2] showed that for the design effects estimated for Armenia and Georgia (roughly 2.5), approximately 85 PSUs with 20 completed interviews per PSU would give sufficient precision. The Armenia team sampled 87 PSUs, and the Georgian team – 85 PSUs. For Azerbaijan, because of the significantly higher design effect, the required number of PSUs would be roughly 115, with 20 completed interviews per PSU.

To account for nonresponse, we assumed that .75 of sampled households would complete an interview, based on previous response rates. Taking into account expected nonresponse, about 30 households would be sampled in each PSU.

---

[2] The formula used to explore the tradeoffs between more clusters and fewer respondents per cluster is the formula for the variance of mean of binary variable sampled under a two-stage design:

$$v(\bar{p}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(p_i - \bar{p})^2 + \frac{1}{n^2(m-1)}\sum_{i=1}^{n} p_i q_i$$

, where n is the number of clusters in the sample, m is the number of respondents in each cluster, p is the proportion of respondents who show a characteristic within cluster i, and q=(1-p). The simulations showed that lowering m did not significantly decrease precision.

Because of the large number of PSUs required for Azerbaijan, it was decided to only sample 90 PSUs, which would still allow for a high degree of precision at the national level, if not within each macro-strata.

Weights

Household Weights

Primary sampling units (PSU) were sampled with probability proportional to size: out of a total N^{psu} PSU's, n^{psu} were sampled. Within each PSU, a fixed number (n^{hh}) of households are sampled. Each PSU had a $m_i^{psu} / \sum m_i^{psu}$ probability of inclusion, where m_i^{psu} is the number of registered voters in the PSU according to the list. If the number of registered voters is an accurate representation of the population of each PSU, then each selected household sample will be self-weighting. The probability of a house being included is $\pi_{ji}^{hh} = n^{hh} / m_i$ where n^{hh} is the number of households sampled in PSU *i*. If π_i^{psu} is the probability that PSU *i* is sampled and π_{ji}^{hh} is the true probability that household *j* is sampled in PSU *i*, then the combined probability that household *j* is sampled is $\pi_{ji} = \pi_i^{psu} \times \pi_{ji}^{hh}$, which is constant across all sampled individuals.

The above weights assume that m^{psu} is directly proportional to the true number of households in the PSU, but in practice, this is probably not true. If an accurate household count was obtained during the sampling process, as it was in Armenia and Georgia, then the weights could be adjusted to reflect the discrepancy between the true number of households and the number of households in the in sampling frame. To do this, one simply uses the following expression: $\pi_{ji} = \pi_i^{psu} \times \pi_{ji}^{est.hh} \times \pi_{ji}^{hh}$, where $\pi_{ji}^{est.hh}$ is probability of selecting a household in the PSU off the voters list and π_{ji}^{hh} is the true probability of selecting a household after an accurate list of households is compiled.

Non-response Adjustment

To account for non-response, the weights for all households who responded in each PSU are adjusted by the inverse of the fraction of sampled households that responded in that PSU. If f_i^{resp} is the fraction of sampled households that responded in PSU *i*, then w_{ji}^{resp}, the non-response weight for household *j* in PSU *i*, is $1 / f_i^{resp}$.

Final Household Weights

To calculate the final household weights, the household weights are multiplied by the non-response adjustment factor: $w_{ji}^{hh} = w_{ji} \times w_{ji}^{resp}$.

These weights are trimmed so that no weight is greater than 4 and rescaled to equal the total number of respondents.

Individual Weights

For individual level data, two sets of weights were constructed: the first account for variation in household size, and the second set are post-stratification weights, that adjust the weights to match sex and age demographic data.

Within Household Selection Weight

Since the number of eligible adults varies from one household to another, the random selection of a single adult introduces inequality into respondents' selection probability. To account for this, a "within household selection weight" is used. The within household weight (w_{ijz}^{within}) for individual *z* in household *j* in PSU z is equal to the size of the household.

Post-Stratification Weights

The individual weights are further adjusted to match census age and sex data. For sex data, for each observation a sex weight is computed that is the fraction of the population belonging to either sex divided by the fraction of the sample of that sex. The sex adjustment factor is thus $w_z^{sex} = F^{sex} / f^{sex}$ for respondent *z*, where F^{sex} is the sex share in the population according to the most recent census figures and f^{sex} is the sex share in the sample. Two adjustment factors are used: one for male respondents (w_z^{males}) and for female respondents (w_z^{female}).

For age[3], respondents' ages are divided into 7 age classes: between 18 and 29, between 30 and 39, between 40 and 49, between 50 and 59, between 60 and 69, between 70 and 79, and 80 and older. For respondents in each weight class, weight adjustment factors are calculated: $w_z^{age} = F^{age} / f^{age}$ for respondent *z*, where F^{age} is the share in the population of that age class according to the most recent census figures, and f^{age} is the share of that age class in the sample.

Final Individual Weights

The final weights are calculated as follows:

$$w_{ijz}^{ind} = w_{ji}^{hh} \times w_{ijz}^{within} \times w_z^{sex} \times w_z^{age}$$

These weights are trimmed so that no weight is greater than 4 and rescaled to equal the total number of respondents.

[3] For Azerbaijan, census data on age was not available. Thus, the individual weights for Azerbaijan only incorporate a sex adjustment factor.