

The analytical domain, the unit at which we wish the data to be statistically representative, is to be the 34 provinces. In addition to these 34 provincial analytical domains there are 10 urban areas with populations larger than 10,000 households. The survey will also collect data to be representative of these 10 urban domains. There are four in total there will be 44 analytical domains.

There are another 21 provinces that have urban households as classified by the CSO household listing. The urban household populations in these 21 provinces ranges from just 64 households in Kapisa to 7,746 in Parwan. These urban populations are both small in number and not thought to be representative of a lifestyle that is so significantly different from the rural households that these households need to be sampled separately from the rural population. Therefore these urban households will be joined together with the rural population in single sample frame to form a mixed rural-urban population for these provinces.

Experience from other surveys suggests that for poverty estimates a sample number of households ranging from a minimum of 350 to 500 produces a stable mean with a reasonably low standard error. It should also be remembered that often there are estimates made on indicators where not all households in the sample provide an observation (e.g. households with school-age children) hence reducing the sample size further.

The rural provincial household population varies enormously according to CSO's latest data for 34 provinces which consists of data for 28 provinces completely enumerated from the CSO household listing, and a further two provinces partially enumerated.

Collecting representative data with a proportional sample a province level creates a challenge because of the large variation in provincial population from the smallest population being in the province of Nimroz with only 13,931 rural households to Herat with 226,576 rural households. If we sample proportionally for Nimroz with 350 households and use the same sampling fraction for all of a households in other provinces the sample would be prohibitively large, more than 80,000 households.

To come within the budget, the provinces Jawzjan with 50,900 rural households is to be used as the base analytical domain for which we wish to determine the sampling fraction. Therefore if we wish to sample 400 households in Jawzjan and use the same sampling intensity for all other provinces of the population greater than that of Jawzjan.

For those provinces with populations less than Jawzjan, if the number of household sampled with the standard sampling fraction is less than 350 households, at further clusters are added when drawing the sample to ensure at least 350 households for these provinces. The sample, therefore is no longer self weighting. The same applies to seven out of the 10 urban clusters, where additional households and clusters have to be drawn to ensure 29 clusters and 348 households.

For all of those analytical domains for which the sampling intensity was standard, with a sample interval of 1519 households, the sample was drawn from one spreadsheet. Therefore when comparing the predicted an actual number of samples drawn from each province there is a difference

of one cluster either way depending on the random start within the order of households within that province. This is not important as for all of these provinces the number of clusters is greater than those needed to meet the minimum 350 households. Also the survey is no longer self-weighting because of the additional households that have to be drawn in some of the urban and rural-the urban provinces with smaller populations.

For those provinces or districts within provinces where the sample frame is not yet available because the CSO household listing exercise has not been conducted in those areas, it may be possible to collect data for two more provinces during the course of the implementation of the NRVA and draw random samples from this frame, otherwise, alternatives will have to be sought which may include using the aims village list and population estimates from which to draw a random sample. The waiting for these provinces can always be adjusted ex-post when the better population estimates is produced by the household listing exercise.

This sample does not include the migratory kuchi population so far.

Village maps with dwellings listed are available from the recent household listing exercise for the 30 provinces that have been enumerated so far, but for only villages greater than 250 households. A scheme for those villagers selected in the random sample was less than 250 households will have to be elaborated.

It is proposed that on arrival at the village, the number of households is determined during the male community interview. This number can be divided by 12, to create a sampling interval for households within the community, and then enumerators select a household each time they count the sampling interval houses. Advantage of this method as it should evenly spread the sampled households throughout the village.

### Sampling for provinces and districts where CSO household listing sample frame is not available.

The following provinces and districts at the time of implementation of the NRVA 2005 did not have household listing data available.

Daikindi Day Kundi	Kandahar Arghistan
Daikindi Gizab	Kandahar Ghorak
Daikindi Kijran	Kandahar Maruf
Daikindi Shahrstan	Kandahar Minashen
Ghazni Nawa	Kandahar Nesh
Ghazni Zana Khan	Kandahar Reg
Hilmand Baghran	Kandahar Shah Wali Kot
Hilmand Dishu	Kandahar Shorabak
Hilmand Garmser	Paktika Dela Khoshmand
Hilmand Gerishk	Zabul Arghandab
Hilmand Kajaki	Zabul Atghar
Hilmand Khanshen	Zabul Daychopan
Hilmand Lashkar Gah	Zabul Khaki Afghan
Hilmand Marja	Zabul Mizan
Hilmand Musa Qala	Zabul Nawbahar
Hilmand Nad Ali	Zabul Qalat
Hilmand Naw Zad	Zabul Seyouray
Hilmand Nawa	Zabul Shahjoy
Hilmand Reg	Zabul Shahre Safa
Hilmand Sangin	Zabul Shumulzayi
Hilmand Washer	Zabul Shinkay

Having considered several options, the use of FAO livestock census village list with estimated number of households was chosen to construct a temporary sampling frame from

which to draw sample villages. This recognises that the FAO list is unlikely to include the urban areas within these districts. Once the CSO household listing has been completed for these provinces, appropriate weights can be applied to the sample in each of these provinces and districts.

### Kuchi sampling for NRVA 2005

The household listing conducted by CSO has not effectively included the migratory kuchi population to date, hence there is no effective sampling frame for this population. Apparently, this lack of enumeration of the kuchi population includes those that have recently settled. This is exactly the same population that was surveyed during winter/spring 2004 by the National Multi-Sectorial Assessment for Kuchi (NAMAK), I the kuchi that is still nomadic and those that have recently settled since the onset of the last drought period. This remains the best estimate of the current kuchi population. The unit of observation for the survey was the kuchi winter location, where one or more kuchi communities may have been located.

The sample frame for this spring summer survey NRVA 05 was created by constructing the predicted kuchi populations their summer location for which information was collected from the NAMAK survey. The geographic location of these summer kuchi community areas is not precise and the best we can do is construct a sampling frame with kuchi communities (donated by tribe and sub tribe name) and their population by district.

We can apply the same sampling fraction as appropriate for the rural settled and urban populations, and we can effectively select communities in their summer location districts. This then gives us a sample frame with a number of kuchi communities to be selected by summer district. Practically, trying to track down the exact community selected in the sampling frame is apparently close to impossible, and therefore we will have to resort to randomly selecting within the number of communities that can be identified from information in the district.

The accompanying spreadsheets indicates the full sample frame with community name and the summary of the number of kuchi communities by summer location district. The latter will be used as the planning tool for drawing the kuchi sampling NRVA 2005.

The summer sample frame produced households in for general kuchi types in the following table.

Kuchi type	HHs
Ex-long range migratory	16,116
Ex-short range migratory	21,282
Long range migratory	123,792
Short range migratory	78,669
Total	239,859

The random sample drawn produced the number following number of primary sample units by kuchi type. 29 primary sampling units is the minimum needed to produce representative data, therefore it is clear that both long and short range migratory kuchi can confidently be analysed as analytical domains. Whereas the ex-long and ex-short range migratory could form a single analytical domain, i.e. those kuchi that a recently ex-migratory but now settled. This recently settled analytical domain would need 5 extra primary sample units to produce a sample for which statistics with a high level of confidence could be calculated.

Because of the indeterminate nature of identifying different communities and therefore different kuchi types, the figures in the table below may not be reflected in the actual sample. Therefore there will need to be some additional questions for the kuchi communities to identify their kuchi type status, and possible *ad hoc* extra primary sample units collected, depending on the frequency of the kuchi types below being selected as the district level. Selection and district level by kuchi type will not be possible, as it will not be possible to reliably identify the kuchi type of communities from information at the district level.

Kuchi type sample	# PSUs
Ex long range migratory	11
Ex short range migratory	12
Long range migratory	84
Short range migratory	51
Margin	158

The table above indicates the number of primary sampling units selected classified by general kuchi type. This may not be reflected in the final sample, as identifying kuchi communities and kuchi type at district level will likely be impossible. Questions will be included in the mail insurer questionnaire to classify each kuchi primary sampling unit by general kuchi type.