

APPENDIX 1f:
Sampling procedure for the Vietnam
Household Living Standard Survey (VHLSS) 2004

This note describes the sampling procedure used during the 2004 Vietnam Household Living Standard Survey (VHLSS 2004).

1. Background

The VHLSS 2004 consists in reality of two surveys: an income only and an income and expenditure survey. A total of 37,200 households were interviewed for the income only survey and a further 9,300 for the income and expenditure survey. Thus, a total of 46,500 household interviews were carried out in 2004.

To understand the sampling procedure it is useful by way of introduction to distinguish among five different layers or aggregation levels, noting that these layers do not in all cases correspond to units used for other government administrative purposes.

- Vietnam as a whole is the most aggregate, national level.
- The 61 provinces of Vietnam make up the second level. Households from all 61 provinces were interviewed during the survey.
- The commune level is the third layer, and 10,511 communes are located across the 61 provinces of Vietnam. Households from 3,100 of these communes were interviewed (either for the income only or for the income and expenditure survey). No households were interviewed from the remaining 7,411 communes.
- Three Enumeration Areas (EAs) in each of the 3,100 communes selected for interviewing make up the so-called master sample.¹ These 9,000 EAs are the fourth level of aggregation. Households from 3,100 of these 9,000 EAs (located in 3,100 different communes) were interviewed for the VHLSS 2004 in clusters of 15 with one cluster from each EA. No households were interviewed from the remaining 5,900 EAs.
- A total of 46,500 individual households make up the fifth and most disaggregated level.

2. Master Sampling

Sampling related to the master sample was carried out in five steps:

¹ Each commune is partitioned into a varying number of Enumeration Areas (EAs) based on the 1999 census, which lists all Vietnamese households. Each EA contains by design around 100 household. The master sample was designed for household surveys to be carried out from 2002-2010.

- The 61 provinces of Vietnam were each allocated a number of communes (adding up as already noted to a total of 3,100) using as weights the proportion of the square root of each provincial population to the sum of the square roots of all 61 provincial populations.²
- The 61 allocated numbers of communes were split by province in two between rural and urban areas according to the provincial rural and urban population shares.
- In each province, separate lists of all rural and urban communes were made in “geographical order”, moving from the North East to the South West provincial corner, registering as well the accumulated number of households as shown in the example below. The sampling of the rural and urban communes in each province was subsequently carried out by applying the interval method to the 122 (2 x 61) rural-urban provincial lists of households.³ Thus, the first commune was in each case picked randomly, and the relevant intervals were calculated by dividing the total number of rural and urban households in a given province with the relevant number of communes to be selected (as set out in the previous step).
- In each of the 3,100 commune selected above, three EAs were sampled randomly giving the master sample of 9,000 EAs.
- In each EA, 15 households were randomly chosen, giving a total master sample of 139,500 households (3,100 communes x 3 EAs x 15 households).

3. VHLSS 2004 Sampling and Interviewing

The 3,100 EAs (corresponding to 46,500 households in clusters of 15) used for the VHLSS 2004 were drawn from the master sample in such a way that households in all 3,100 communes were interviewed in one of the two types of surveys. Accordingly, in each of the 3,100 communes, one EA out of the three in the master sample was randomly selected.

These 3,100 EAs (from each their commune) were subsequently randomly divided into two groups of respectively 2,350 and 750 enumeration areas. This was done to arrive at the actual

2 The use of the square roots as weights implies that thinly populated provinces are overrepresented in the national sample. Appropriate adjustments therefore have to be made when national results are derived. However, the approach taken does make it possible to derive representative conclusions at the province level.

3 The interval method means that if for example five different areas (communes) have to be selected from a list of 30 the analyst selects one area randomly and then every sixth entry starting with the randomly selected one. This implies in the present example that area 1 and 7 are selected together; as are 2 and 8, 3 and 9 etc.

households, which were interviewed in respectively the income only (9,300) and income and expenditure (37,200) surveys.

The actual interviewing was carried out during 2004 as follows:

- In May, half of the sample of 46,500 households were interviewed. Of which 4650 households (from 310EAs) interviewed income only and 18,600 households (from 1,240 EAs) were interviewed expenditure.
- In September, the remain households of the sample were interviewed.

Sampling Methodology for the Master Sample - VLSS2004



