CHAPTER
Two

Methodology

2.1 Coverage of the Labour Force Survey

The LFS was a nationwide survey covering household population in all the nine provinces both rural and urban areas. The survey excluded institutional populations such as those in hospitals, Barracks or refugee camps.

2.2 Questionnaire

The questionnaire used to collect child labour data was part of the labour force survey questionnaire which also included other subjects as Time-Use (TU) and Information and Communication Technology (ICT). The questionnaire was adapted through a series of technical meetings with stakeholders such as the Ministry of Labour, International Labour Organization (ILO) and other related agencies.

The questionnaire was arranged in nine sections namely:

1 Background Characteristics
2 Education and School Attendance
3 Economic Activity
4 Unemployment
5 Employment
6 Health and Safety Issues of persons five years and Above
7 Child Labour
8 Time Use
9 Information and Communications Technology (ICT)

2.3 Sampling Design and Implementation

The sample was designed to allow separate estimates for the nation as a whole, and rural and urban areas. The sample design also allowed for indicators to be estimated for each of the nine provinces.
A representative probability sample of 8000 households was selected in two stages. In the first stage 320 clusters were selected from a list or frame of enumeration areas compiled from the 2000 census of population and housing. A cluster is the primary sampling unit which is equivalent to a Standard Enumeration Area (SEA). In the second stage a complete listing of households in the selected clusters was carried out. Households were then selected systematically for participation in the interviews.

2.3.1 Sampling Frame and Stratification

The country, Zambia is administratively, divided into nine provinces. Each province is in turn subdivided into districts. For statistical purposes each district is subdivided into Census Supervisory Areas (CSAs) and these are in turn subdivided into Standard Enumeration Areas (SEAs). The 1998-2000 mapping exercise in preparation for the 2000 census of population and housing, demarcated the CSAs within in wards, wards within constituencies and constituencies within districts. In total, Zambia has 72 districts, 150 constituencies, 1,289 wards, about 4,400 CSAs and about 17,000 SEAs. The listing of SEAs has information on number of households and the population. The number of households will be used as a measure of size for selecting primary sampling units (PSU). Therefore, the sample frame of this survey is the list of SEAs developed from the 2000 Population Census.

The SEAs are also stratified by urban and rural strata.

2.3.2 Sample allocation and selection

The total sample of 8,000 households was first allocated between rural, urban and the provincial domains in proportion to the population of each domain according to the 2000 census results. The proportional allocation does not however allow for reliable estimates for smaller domains. Adjustments to the proportional allocation of the sample were made to allow for reasonable comparison to be achieved between strata or domains. Therefore disproportionate allocation was adopted, for the purpose of maximizing the precision of survey estimates**.** The disproportionate allocation is based on the square root method designed by Leslie Kish. The sample was then selected using a stratified two-stage cluster design.
The distribution of sample clusters and households based on the disproportionate allocation is given in the table below.

Table 2.1: Sample Allocation of Clusters According to Province and by Type of Residence

Province	Allocation of clusters		
	Rural	Urban	Total
Central	24	10	34
Copperbelt	11	31	42
Eastern	35	3	38
Luapula	28	4	32
Lusaka	9	33	42
Northern	35	5	40
North-Western	23	3	26
Southern	28	8	36
Western	26	4	30
Total	219	101	320

After the households were allocated to the different strata, the number of clusters to be selected was calculated based on an average cluster take of 25 completed interviews of all households. Clusters were selected systematically with probability proportional to the number of households.

Sélection of clusters

The procedure for selecting clusters (i.e. SEAs) in each stratum involved:

(i) Calculating the sampling interval, I, for each stratum

$$I_h = \frac{\sum_{i=1}^{N_h} M_{hi}}{a_h}$$

where M_{hi} is the number of households in SEA (or cluster) i and stratum h, $\sum_{i=1}^{N_h} M_{hi}$ is the size of the stratum (total number of households in the stratum according to the 2000 census) and a is the number of clusters (SEAs) to be selected in the stratum.

(ii) Calculating the cumulated size of each SEA.

(iii) Calculating the sampling numbers

R, $R+I$, $R+2I$, ..., $R + (a\text{-}1)I$, where R is a random number between 1 and I.

(iv) Comparing each sampling number with the cumulated sizes of the SEAs.

The first SEA (or cluster) whose cumulated size is equal to or greater than the random number (ii) was selected. The next SEA to be selected is the one with cumulated size equal to or greater than $R+I$. Each of the rest of the SEAs were selected using the same procedure, making sure to add I at each subsequent selection.

Selection of households

A frame of households was determined by listing all the households in all the selected SEAs. During listing a couple of questions were asked in order to group households into three categories: those with at least one paid child worker, those with at least one unpaid child worker and those that had no working children at all.

The total number of households to be selected per cluster was allocated between the different categories in proportion to the households found in each category. Once the number of households to be selected in each category was calculated, the following steps were used to select the households from each category:

1. Calculate the sampling interval for each category

$$I_b = \frac{B}{b}$$

where B is the number of households listed in the selected SEA and b is the number of households to be selected in the selected SEA.

1 Generating a random number (R) between 1 and the Interval I; the first selection was hence R

2 Adding the interval to the random number to get the next selection

4. Adding the interval repeatedly until the desired sample size was achieved.

2.4 Pre-test

The Pre-test for the child labour Survey was conducted in October 2005.The objective of the Pre-test was to test the adequacy of the survey instruments and also served as an opportunity to train Trainers for the main survey. The participants in the Pre-test included the survey implementation team members and those who were to train in the main training.
The Pre-test exercise consisted of two parts. The first part which took one week comprised of training following a classroom set-up and the second part was meant for field work. The training included role plays and demonstration interviews using local language. Field practice lasted for three days. The participants for the pre-test met after the exercise for a debriefing and shared experiences which formed a basis for finagling the instruments.

2.5 Training of Interviewers and Supervisors

About 170 persons were recruited by the Central Statistical Office to serve as enumerators and supervisors. They all participated in the main training which begun on … 2005 in the different provinces. Staff from the CSO and Ministry of Labour who were part of the technical team led the training which was conducted for two weeks. Training was guided by the enumerators manual that was prepared as part of the survey instruments. The method of training included having enumerators read from the manual and trainers lecturing on different topics in line with what the manual prescribed. Other training modes included class demonstrations (front of class interviews) and interviews in small groups. The method was exactly what was done during the Pre-test. Two days was allocated for field practice.

2.6 Fieldwork

At least three enumerators were assigned to one supervisor and they formed up a team. Every province had more than three teams. Depending on the availability of transport and other logistics associated with a province, field work was accomplished by team work (working in one cluster per time as a team) or assigning every enumerator a cluster to work in.
Data collection was conducted between November and December 2005. The Central Statistical Office coordinated the supervision of fieldwork. Trainers visited the field teams during the initial implementation of fieldwork. The provincial /statisticians who also attended training monitored the quality of data in the field. There was close contact between the Field teams and Headquarters which was maintained through out fieldwork.

2.7 Data Processing

During training of field staff the data processing team also sat in the training to familiarize themselves with the questionnaire. After field work begun a training session was undertaken in each process on editing and data entry. Completed questionnaires were submitted to provincial offices for editing and data entry. Basically four stages were involved in data processing namely, manual editing and coding, data entry, data cleaning and tabulation.

2.8 Response rates and weighting

Table2.2 shows response rates for the Labour Force Survey. A total of 8,000 households were selected in the sample of which 7,886 were occupied and therefore were eligible for interviews. The difference was mainly due to non availability of household members that could respond to the questionnaire. The overall response rate for household interviews was 98.5 percent.

Table 2.2: Household Interview Response Rates

Result	Rural	Urban	Total
Number of Households Selected	5,475	2,525	8,000
Number of Households Occupied	5,440	2,446	7,886
Household Response Rates	99.4	96.9	98.5

2.8.1 Estimation Procedure

Weights

Due to the non-proportional allocation of the sample to the different strata, sampling weights were required to ensure actual representative ness of the sample at national level. The sampling probabilities at first-stage selection of SEAs and probabilities of selecting the households were used to calculate the weights. The weights of the sample are equal to the inverse of the probability of selection.

The probability of selecting cluster *i* was calculated as

$$P_{hi} = \frac{a_h M_{hi}}{\sum_{i=1}^{M_h} M_{hi}}$$

The weight or boosting factor is, thus, given as

$$w_{hi} = \frac{1}{P_{hi}}$$

where: p_{hi} is the first stage sampling probability of (SEA), a_h is the number of SEAs selected in stratum *h*, M_{hi} is the size (households according to the Census frame) of the i^{th} SEA in stratum *h*, and ΣM_{hi} is the total size of stratum *h*.

The selection probability of the household was calculated as:

$$p_h = \frac{n_h}{N_h}$$

where n_h =the number of households selected from stratum *h*, N_h=the total number of households in stratum *h*.

Let y_{hij} be an observation on variable *Y* for the j^{th} household in the i^{th} SEA of the h^{th} stratum. Then the estimated Total for the h-th stratum is:

$$y_h = \sum_{i=1}^{a_h} \sum_{j=1}^{n_h} w_{hi} y_{hij}$$

where, y_h is the estimated total for the h-th stratum., w_{hi} is the weight for the j^{th} household in the i-th SEA of the h-th stratum, i=1-a_h is the number of selected clusters in the stratum, j=1-n_h is the number of sample households in the stratum. The national estimate is given by:

$$y = \sum_{h=1}^{H} y_h$$

where, *y* is the national estimate, h=1, …, *H* is the total number of strata. For this survey, *H* = 18 (the nine provinces by rural and urban taken as a separate domains).

2.9 Quality Control

In order to ensure reliability and credibility of data collected some quality control measures designed for the survey included formation of a technical team which had members from CSO, Ministry of Labour and other relevant stake holders. These were involved in the planning and implementation of the survey. The trainers monitored field work through out and the supervisors remained with their assigned teams to edit the work on daily basis. The review of generated tables from the survey also involved the stakeholders that took part in the planning.