APPENDIX A: SAMPLE DESIGN FOR PAKISTAN SOCIAL AND LIVING STANDARDS MEASUREMENT SURVEY, 2005-06

Objectives:

The data generated though PSLM Survey will be used to assist the government in formulating the poverty reduction strategy in the overall context of MDGs. The indicators will be developed at National/Provincial level in the following sectors.

1. Education
2. Health
3. Water Supply & Sanitation.
4. Population Welfare
5. Income & Expenditure

Universe:

The universe of this survey consists of all urban and rural areas of all four provinces. Military restricted and protected areas of NWFP have been excluded from the scope of the survey.

SAMPLING FRAME

Urban area:

FBS has developed its own urban area frame. All urban areas comprising cities/ towns have been divided into small compact areas known as enumeration blocks (E.Bs) identifiable through map. Each enumeration block comprises about 200-250 households and categorized into low, middle and high-income group, keeping in view the socio economic status of the majority of households. Urban area sampling frame consists of 26698 enumeration blocks has been updated in 2003.

Rural area:

With regard to the rural areas, the lists of villages/mouzas/dehs according to Population Census, 1998 have been used as sampling frame. In this frame, each village/mouza/deh is identifiable by its Name, Had Bast Number, Cadastral map etc. This frame is comprised of 50590 villages/mouzas.

The numbers of enumeration blocks in urban and mouzas/dehs/villages in rural areas of the country are as under:

NO. OF ENUMERATION BLOCKS AND VILLAGES AS PER SAMPLING FRAME

Province	Number of E. Blocks	Number of Villages
Punjab	14,549	25,875
Sindh	9,025	5,871
NWFP	1,913	7,337
Balochistan	613	6,557
A.J.K	210	1,654
Northern Area	64	566
FATA		2,596
Islamabad	324	132
Total	**26,698**	**50,588**

STRATIFICATION PLAN

Urban Areas:

Large sized cities having population five lacs and above have been treated as independent stratum. Each of these cities has further been sub-stratified into low, middle and high income groups. The remaining cities/towns within each defunct administrative division have been grouped together to constitute an independent stratum.

Rural Areas:

The entire rural domain of a district for Punjab, Sindh and NWFP provinces has been considered as independent stratum, whereas in Balochistan province defunct administrative division has been treated as stratum

Sample Size and its Allocation:

To determine optimum sample size for this survey, analytical studies based on the results of Pakistan Demographic Survey, Labour Force and Pakistan Integrated Households Sample Survey were undertaken. Keeping in view the variability exist within the population for the characteristics for which estimates are to be prepared, population distribution, level of estimates and field resources available a sample size of 15453 households enumerated from 1109 sample PSUs (531 from urban and 578 from rural areas) has been considered sufficient to produce reliable estimates in respect of all provinces. The distribution plan of PSUs and SSUs by province and region is as under:-

PROFILE OF THE PSLM SAMPLE 2005-06

PROVINCE	2005-06 PSLM		
	URBAN	RURAL	TOTAL
PSUs:			
Punjab	240	244	484
Sindh	140	132	272
NWFP	88	119	207
Balochistan	63	83	146
Overall	**531**	**578**	**1109**
SSUs/Households			
Punjab	2790	3892	6682
Sindh	1666	2107	3773
NWFP	1049	1901	2950
Balochistan	735	1313	2048
Overall	**6240**	**9214**	**15453**

Sample Design: A two-stage stratified sample design has been adopted for this survey.

Selection of primary sampling Units (PSUs):

Enumeration blocks in the urban domain and mouzas/dehs/villages in rural domain have been taken as primary sampling units (PSUs). In urban domain sample PSUs from each stratum have been selected by probability proportional to size (PPS) method of sampling scheme using households in each block as measure of size (MOS). Similarly in rural areas, population of each village has taken as MOS for selection of sample villages using probability proportional to size method of selection.

Selection of Secondary Sampling Units (SSUs):

Households within each sample Primary Sampling Unit (PSU) have been considered as Secondary Sampling Units (SSUs). 16 and 12 households have been selected from each sample village and enumeration block respectively by systematic sampling scheme with a random start.

Estimation Procedures:

Detail of estimation procedures for estimates and their variances is attached as Annexure – I

Annexure-I

Estimation Procedure:

ESTIMATION PROCEDURE ADOPTED FOR PSLM SURVEY

NOTATIONS:

N_h = Total number of Primary Sampling Units (PSUs) in the hth stratum of a province.

n_h = Total number of sample PSUs in the hth stratum of a province.

M_{hi} = Total number of Secondary Sampling Units (SSUs) in the ith sample PSU of hth stratum of a province.

m_{hi} = Number of sample SSUs in the ith sample PSU of hth stratum of a province.

P_{hi} = Assigned probability of selection of ith PSU of the hth stratum of a province.

y_{hij} = Value of any characteristic y of jth SSU within ith PSU of hth stratum of a province.

x_{hij} = Value of any characteristic x of jth SSU within ithPSU of hth stratum of a province with whose respect proportion is required.

(i): ESTIMATION FORMULAE FOR TOTALS AND THEIR VARIANCES

$$N = \sum_{h=1}^{L} N_h$$

$$n = \sum_{h=1}^{L} n_h$$

$$\hat{Y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{p_{hi}}$$

OR

$$\hat{Y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{1}{p_{hi}} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} y_{hij}$$

$$\hat{Y} = \sum_{h=1}^{L} \hat{Y}_h = \sum_{h=1}^{L} \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{p_{hi}}$$

For X, another variable of interest, we have

$$\hat{X}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\hat{X}_{hi}}{P_{hi}} = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{1}{P_{hi}} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} x_{hij}$$

$$\hat{X} = \sum_{h=1}^{L} \hat{X}_h = \sum_{h=1}^{L} \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\hat{X}_{hi}}{p_{hi}}$$

$$\hat{R} = \frac{\hat{Y}}{\hat{X}}$$

$$v(\hat{y}_h) = \frac{1}{n_h} s^2_{ht} = \frac{1}{n_h (n_h - 1)} \left(\sum_{i=1}^{n_h} \frac{\hat{Y}^2_{hi}}{P^2_{hi}} - \frac{(\sum_{i=1}^{n_h} \frac{\hat{y}_{hi}}{P_{hi}})^2}{n_h} \right)$$

$$v(\hat{Y}) = \sum_{h=1}^{L} \frac{1}{n_h} s^2_{ht} = \sum_{h=1}^{L} \frac{1}{n_h (n_h - 1)} \left(\sum_{i=1}^{n_h} \frac{\hat{Y}^2_{hi}}{P^2_{hi}} - \frac{(\sum_{i=1}^{n_h} \frac{\hat{y}_{hi}}{P_{hi}})^2}{n_h} \right)$$

(ii): FORMULA FOR RATIO ESTIMATES

$$r = \frac{\hat{Y}}{\hat{X}}$$

where $Y^{\wedge}$ and $X^{\wedge}$ can be estimated by equations under item (i) given above.

$$Rel\ V(r) = \frac{1}{\hat{X}^2} \sum_{h=1}^{L} \frac{1}{n_h} s^2_{hb} + \frac{1}{\hat{x}^2} \sum_{h=1}^{L} \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{M^2_{hi}}{p^2_{hi} m_{hi}} \frac{(M_{hi} - m_{hi})}{M_{hi}} s^2_{hw}$$

where

$$s^2_{hb} = s^2_{ht} - s^2_{hw}$$

$$s^2_{ht} = s^2_{hy} + r^2 s^2_{hx} - 2r\, s_{hxy}$$

$$s^2_{hx} = \frac{1}{(n_h - 1)} \left[\sum_{i=1}^{n_h} \frac{\hat{x}^2_{hi}}{p^2_{hi}} - \frac{\left(\sum_{i=1}^{n_h} \frac{\hat{x}_{hi}}{p_{hi}} \right)^2}{n_h} \right]$$

$$s^2_{hy} = \frac{1}{(n_h - 1)}\left[\sum_{i=1}^{n_h}\frac{\hat{y}^2_{hi}}{p^2_{hi}} - \frac{\left(\sum_{i=1}^{n_h}\frac{\hat{y}_{hi}}{p_{hi}}\right)^2}{n_h}\right]$$

$$s_{hxy} = \frac{1}{n_h - 1}\left[\sum_{i=1}^{n_h}\left(\frac{\hat{X}_{hi}}{p_{hi}}\frac{\hat{y}_{hi}}{p_{hi}}\right) - \frac{\left(\sum_{i=1}^{n_h}\frac{\hat{X}_{hi}}{P_{hi}}\right)\left(\sum_{i=1}^{n_h}\frac{\hat{y}_{hi}}{P_{hi}}\right)}{n_h}\right]$$

$$s^2_{hw} = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}\frac{1}{p^2_{hi}}\frac{M^2_{hi}}{m_{hi}}\frac{(M_{hi} - m_{hi})}{M_{hi}}s^2_{hi}$$

and

$s^2_{hi} = s^2_{hiy} + r^2 s^2_{hix} - 2r\, s_{hixy}$

$$s^2_{hiy} = \frac{1}{(m_{hi} - 1)}\left[\sum_{j=1}^{m_{hi}} y^2_{hij} - \frac{\left(\sum_{j=1}^{m_{hi}} y_{hij}\right)^2}{m_{hi}}\right]$$

$$s^2_{hix} = \frac{1}{(m_{hi} - 1)}\left[\sum_{j=1}^{m_{hi}} x^2_{hij} - \frac{\left(\sum_{j=1}^{m_{hi}} x_{hij}\right)^2}{m_{hi}}\right]$$

$$s^2_{hixy} = \frac{1}{(m_{hi} - 1)}\left[\sum_{j=1}^{m_{hi}} x_{hij}\, y_{hij} - \frac{\left(\sum_{j=1}^{m_{hi}} x_{hij}\sum_{j=1}^{m_{hi}} y_{hij}\right)}{m_{hi}}\right]$$