# LABOUR FORCE SURVEY – TURKEY

## 1. General Information

Turkey has planned to pass to implement Labour Force Survey continously. For this reason during year 2013, two surveys (existing LFS and continous LFS) were carried out at the same time to see if there are changes resulting from methodology and implementation of continous LFS.

For existing LFS;

- → *2 - 2 - 2 design is used*
- → *Around 42.000 households per quarter*
- → *353 clusters (15 hh selected from each) for each subsample*
- → *Regulation 577/98 Article 3(1) is followed for sample size calculation*
- → *1st week of the month is the reference week (12 per year)*

For continous LFS;

- → *2 - 2 - 2 design is used*
- → *Around 44.000 households per quarter*
- → *550 clusters (10 hh selected from each) for each subsample*
- → *Regulation 577/98 Article 3(1) and 3(1) is followed for sample size calculation*
- → *Each week is reference week (52 per year)*

## 2. Sample Design

2in – 2out – 2in rotational design and 8 Subsamples are used for LFS. Each quarter 1 subsample is going out of the sample while the new one enters. By this sample design, %50 overlap is aimed between consecutive quarters and again %50 overlap between the same quarter of consecutive years.

| | 2013 | | | | 2014 | | | | 2015 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Subsample** | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** |
| **1** | A | | | | | | | | | | | |
| **2** | **A** | **A** | | | | | | | | | | |
| **3** | B | A | A | | | | | | | | | |
| **4** | **B** | **B** | A | A | | | | | | | | |
| **5** | A | B | B | A | A | | | | | | | |
| **6** | **A** | **A** | B | B | A | A | | | | | | |
| **7** | B | A | A | B | B | A | A | | | | | |
| **8** | **B** | **B** | A | A | B | B | A | A | | | | |
| **9** | | B | B | A | **A** | B | B | A | **A** | | | |
| **10** | | | B | B | **A** | A | B | B | **A** | A | | |
| **11** | | | | B | **B** | A | A | B | **B** | A | A | |
| **12** | | | | | **B** | B | A | A | **B** | B | A | A |
| **13** | | | | | | B | B | A | A | B | B | A |
| **14** | | | | | | | B | B | A | A | B | B |
| **15** | | | | | | | | B | B | A | A | B |
| **16** | | | | | | | | | B | B | A | A |
| **17** | | | | | | | | | | B | B | A |
| **18** | | | | | | | | | | | B | B |
| **19** | | | | | | | | | | | | B |

## 3. For the calculation of sample size

- precision
- level of confidence interval
- estimated or known proportion of the population in specified group
- predicted coverage rate or prevalence for the specified indicator
- design effect
- average household size and
- adjustment for potential loss of sample household due to nonresponse is used.

Mathematical equation for calculating the sample size is the following:

$$n = \frac{t^2.p.q.deff}{d^2.r.n} H$$

In the formula;

t : factor to achieve %95 percent confidence interval

p : proportion of the characteristic

q : (1-p)

d : degree of precision

r : the proportion of the total population in which the indicator $p_i$ calculated.

n : average household size

H : number of reporting domains

After several alternatives were studied, a kind of compromise allocation has been applied for the allocation of the sample size to strata. This allocation method gives result between proportional and equal allocation and ensures a minimum sample size for each domain.

**Firstly;** Sample size calculation has started at NUTS1 * Urban – Rural

**Secondly**; The results from the first step are used to calculate sample sizes at NUTS2 Total level

**Finally;** Sample sizes are distributed to NUTS3 (provinces) and initial sample sizes were gathered before checking for EUROSTAT requirements.

After we calculated initial sample sizes, we checked if these sample sizes satisfy the requirements (article 3(1) and article 3(2)) stated in "Doc.: Eurostat/F2/LAMAS/38/10" by using household sizes, CV values and nonresponse rates at NUTS2 level and updated the sizes.

After final sample sizes are calculated, sample clusters are distributed uniformly among reference weeks (13 weeks per quarter) at nuts3 * urban – rural level.

**4. Weighting:**

Following procedures are going to be used for weighting of LFS in Turkey:

- Design weights, calculated depending on the selection probabilities of the enumeration areas (clusters) in related domains.
- Non response adjustment factors, calculated at cluster level.
- Integrative calibration of population at Age * Gender and Nuts 2 * Urban –Rural level according to the projected populations. This step will allow consistent results with total population distribution in terms of age, gender, nuts2 and urban-rural.
- Trimming of weights for reducing extreme variations in the sampling weights.

**4.1. Design weights**

Design weights are inversely proportional to the overall selection probabilities. At first stage, clusters were selected from domains. After sample clusters were selected, we proceed with the selection of households from each cluster. 10 households are selected from each cluster.

f1i : selection probability of clusters from i'th domain. (each domain has different number of clusters and different number of clusters are selected from different domains)

f1i = # of selected clusters / # of total clusters in domain

f2 : selection probability of households from each cluster.

f2 = 10 / (cluster size)

f0 : selection probability of a household

f0 = (1/f1i)*(1/f2)

base weight (w) : 1 / f0

**4.2. Weighting for Nonresponse**

In order to reduce the effect of nonresponse, adjustment factor related to nonresponse rate is calculated in design domains at cluster level. This factor is inversely proportional to response rate for households. In the survey, nonresponse adjustment is made at household level in each cluster.

***Response rate at household level :***

$$R_i = \frac{\text{no of interwieved households in ea i}}{\text{no of eligible households in ea i}}$$

Nonresponse adjustment factor is inversely proportional to response rate. Again it can be scaled to have an average value of 1.0

$$w_i = \frac{\overline{R}}{R_i}$$

The numerator, being the overall response rate, is defined below in terms of the number of interviewed cases in cluster.

$$\overline{R} = \frac{n_i}{\sum \frac{n_i}{R_i}}$$

Also no calculation is made concerning individual nonresponse.

**4.3. Integrative calibration**

Calibration is applied to ensure the weighted sample sums of specified control variables equal to the known population totals.

One of the practical calibration methods is “iterative proportional fitting or raking”. In this approach the weighted sample distribution is adjusted to the external controls simultaneously with an iterative process.

For Labour Force Survey (LFS), integrative calibration procedure, based on iterative proportional fitting, is going to be applied so that the weight given to each member within the household will be equal to the household weight and consistent results with population projection will be achieved in terms of sex, age and nuts2 * urban - rural characteristics.

***The formulas of integrative calibration:***
In the integrative calibration, the weighted proportion of individuals in the group of interest is calculated below:

$$p_i = \frac{\sum Dw_k}{\sum w_k}$$

In the formula :
D : (0-1) variable having the value 1 for the induviduals in the group of interest, and 0 otherwise.
$w_k$ : design weight of the kth individual.
$p_i$ : the weighted proportion of individuals in the group of interest
$P_i$ : proportion of indviduals in the group of interest from the external source
i : index of the group of interest

In order to calculate the weights of integrative calibration, a variable $D_h$ is defined in the following way :

If there is at least one individual having the value D=1 in household h, then other individuals in the same household is assigned the $D_h$ value equal to 1. Accordingly, if no individuals in the $h^{th}$ household is in the group of interest (all individuals in the $h^{th}$ household have the value D=0), then the value of $D_h$ is equal to 0.

The weight of the $h^{th}$ household is calculated in two cases by depending on the value of $D_h$:

If $D_h$ =1 then

$$w_h(I) = \frac{P_i}{p_i} w_h(I)old$$

If $D_h$=0 then

$$w_h(II) = \frac{1 - a_i\left(\frac{P_i}{p_i}\right)}{1 - a_i} w_h(II)old$$

where $a_i$ is equal to the weighted proportion of $D_h$ values in the specified group i.

This process is repeated iteratively until the sample distribution is fitted to population distribution.

**4.4. Trimming**

Control of extreme values and large variations in weights is desirable at the stage of adjusting the weights, after nonresponse adjustment and calibration. A common approach for this control is to trim extreme values.

Just after the calculation of non-response and calibration weights, any computed weight outside the following limits is recoded to the boundary of these limits:

$$\frac{1}{L} \le \frac{w_h}{\overline{w}} \le L$$

In the formula $\overline{w}$ is the mean value of household calibrated weight -at the same time the calibrated weight of individuals within the household- and L is some appropriate upper bound for the adjustment in weights. The value of L is going to be taken as 3 as being a resonable value for determining the boundaries.