

APPENDIX A METHODOLOGY

1. Sample Design

1.1 Establishments with 1 – 10 persons engaged

The stratified systematic sampling was adopted for the survey. Bangkok and each province was constituted a stratum. The sampling units were establishments.

1.1.1 Stratification

Bangkok and each province was constituted a stratum. There were altogether 77 strata. Each stratum was classified by TSIC at activities level into 309 sub-stratum.

1.1.2 Selection of Sampling Unit

The sample selection of establishments was done by systematic sampling and performed separately and independently in each sub-stratum. The total sample were 140,766 from 1,685,335 establishments.

1.2 Establishments with 11 persons engaged and over

All the establishments with 11 persons engaged and over were completely enumerated. There were 51,010 establishments.

In summary, the total number of sample establishments selected for enumeration were 191,776 classified by region and size as follows :

region	Total	1 – 10 persons engaged	11 persons engaged and over
Bangkok	27,056	8,263	18,793
Vicinity	16,514	11,494	5,020
Central	40,459	32,011	8,448
North	36,471	30,602	5,869
Northeast	40,309	33,404	6,905
South	30,967	24,992	5,975
Whole kingdom	191,776	140,766	51,010

2. Estimation method

The survey results were presented at regional level namely Bangkok, Vicinity (Samut Prakan, Nonthaburi, Pathum Thani, Nakhon Pathom and Samut Sakhon), the North, the Northeast and the South. All the establishments were classified by TSIC at activities level and divided into 6 sizes according to number of persons engaged as follows:

Size of establishment	1	2	3	4	5	6
Number of persons engaged	1 – 15	16 – 25	26 – 30	31 – 50	51 – 200	> 200

Let,

$q = 1, 2, 3, \dots, n_{hijklmp}$ (sample establishment)

$p = 1, 2, 3, \dots, 12$ (size of establishment)

$m = 1, 2, 3, \dots, 309$ (activity)

$l = 1, 2, 3, \dots, 124$ (class)

$k = 1, 2, 3, \dots, 75$ (group)

$j = 1, 2, 3, \dots, 27$ (division)

$i = 1, 2, 3, \dots, A_h$ (province)

$h = 1, 2, 3, 4, 5, 6$ (region)

2.1 The estimation for establishment with 1 – 10 persons engaged

The estimation total number of characteristics X of the establishments with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region was based on the formula:

$$\hat{X}'_{hijklm} = \sum_{i=1}^{A_h} \hat{X}'_{hijklm} \dots\dots\dots (1)$$

wherer

\hat{X}'_{hijklm} is the estimate of the total number of characteristics X of the establishments with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region. Where

$$\hat{X}'_{hijklm} = w'_{hijklm} \sum_{p=1}^2 \sum_{q=1}^{n_{hijklmp}} x_{hijklmpq}$$

$x_{hijklmpq}$ is the value of characteristics X for the q^{th} sample establishment, p^{th} size of establishment, m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

w'_{hijklm} is the weighting factor of the establishment with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region. Where

$$w'_{hijklm} = \frac{N'_{hijklm}}{n'_{hijklm}}$$

N'_{hijklm} is the total number of the establishments with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

n'_{hijklm} is the total number of the sample establishments with 1 – 10 persons

engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

A_h is the total provinces in the h^{th} region, where $\sum_{h=1}^6 A_h = 77$

2.2 The estimation for establishment with 11 persons engaged and over

The estimation total number of characteristics X of the establishments with 11 persons engaged and over ($p = 3, 4, 5, \dots, 12$) for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region was based on the formula:

$$\hat{X}_{hijklmp} = \sum_{i=1}^{A_h} \hat{X}_{hijklmp} \dots\dots\dots (2)$$

where

$\hat{X}_{hijklmp}$ is the estimate of the total number of characteristics X of the establishments for the p^{th} size of establishment, m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region. Where

$$\hat{X}_{hijklmp} = \frac{N_{hijklmp}}{N'_{hijklmp}} \sum_{q=1}^{n_{hijklmp}} X_{hijklmpq}$$

$X_{hijklmpq}$ is the value of characteristics X for the q^{th} sample establishment, p^{th} size of establishment, m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

$N_{hijklmp}$ is the total number of the establishments for the p^{th} size of establishment, m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

$N'_{hijklmp}$ is the total number of the responding establishments for the p^{th} size of establishment, m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region.

2.3. The estimation of coefficient of variation of the estimate of the total number

1) The estimation of the variance of the estimate of the total number of characteristics X of the establishments with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, i^{th} province, h^{th} region was based on the formula:

$$\hat{V}(\hat{X}'_{hijklm}) = N_{hijklm} (N_{hijklm} - n_{hijklm}) \frac{s_{hijklm}^2}{n_{hijklm}}$$

where

$$s_{hijklm}^2 = \frac{1}{n_{hijklm} - 1} \left[\sum_{q=1}^{n_{hijklm}} x_{hijklmq}^2 - \frac{\left(\sum_{q=1}^{n_{hijklm}} x_{hijklmq} \right)^2}{n_{hijklm}} \right]$$

2) The estimation of the coefficient of variation of the estimate of the total number of characteristics X of the establishments with 1 – 10 persons engaged for the m^{th} activity, l^{th} class, k^{th} group, j^{th} division, h^{th} region was based on the formula:

$$C.V. \cdot h_{ijklm} = \frac{\sqrt{\sum_{i=1}^{A_h} \hat{V}(\hat{X}'_{hijklm})}}{\sum_{i=1}^{A_h} \hat{X}'_{hijklm}}$$

3. Data collection

The interviewing method was employed in data collection. The enumerators who are permanent and temporary staff of the National Statistical Office were sent out to interview the owners or the entrepreneurs of the sampled business establishments during May – August 2012.

4. Errors of the data

Data presented in this report might be subject to sampling and non-sampling errors. For instance, errors from the imputation for missing values and non-response, intentional misreporting and errors arising at coding and data entry stages. However, the NSO tried its best to minimize such errors, thus the data should be used with appropriate cautions.

5. In round figures

The summation of each amount may not equal to the total due to rounding.