

**Vietnam Employer Survey**  
**Construction of Sample Weights**

## 1 Sample Design

The Viet Nam Employer Survey used a stratified random sample design.

The Sample Frame used for the construction of sample weights was the worksheet 'Firms in scope' in the Excel file 'Firms.original'.

The sample frame was stratified into 3 Firm Size categories. A sample of 594 firms was selected from the sample frame. Table 1 shows the number of firms in the sample frame, i.e., Population Size, and the corresponding final sample size by Firm Size stratum.

<b>Table 1: Population and Sample Sizes by Stratum</b>		
<b>Stratum – Firm Size (# of employees)</b>	<b>Population Size</b>	<b>Final Sample Size</b>
<b>1 to 10</b>	16,519	189
<b>11 to 50</b>	8,631	200
<b>51 or more</b>	4,764	205
<b>Total</b>	<b>29,914</b>	<b>594</b>

## 2 Survey Results

Two final sample files with the final survey results were used for the weighting of the Viet Nam Employer Survey:

- 1) [Employer\\_sample\\_594 firms-Jul 24](#) – This sample file contains the final response codes for the 594 sampled firms. The final response codes are,
  - 1 – Complete
  - 2 – Ineligible
  - 3 – Refusal
  - 4 – Non-response
- 2) [Employer\\_survey\\_dataset\\_July24](#) – This sample file contains the collected questionnaire data for 330 sampled firms with a response code '1- Complete'.

'A final sample of 330 completed questionnaires was obtained from the sample of 594 firms. Table 2 shows for each stratum the response results from the sample of 594 firms. The column 'Non-response' includes the combined stratum sample counts for the response codes '3 – Refusal' and '4 – Non-response'.

<b>Table 2: Response by Stratum</b>					
<b>Stratum – Firm Size (# of employees)</b>	<b>Population Size</b>	<b>Final Sample Size</b>			
		<b>Complete</b>	<b>Ineligibles</b>	<b>Non-response</b>	<b>Total</b>
<b>1 to 10</b>	16,519	87	41	61	189
<b>11 to 50</b>	8,631	114	25	61	200
<b>51 or more</b>	4,764	129	11	65	205
<b>Total</b>	<b>29,914</b>	<b>330</b>	<b>77</b>	<b>187</b>	<b>594</b>

### 3 Variables Added to Sample Dataset

During the weighting process several variables were added to the sample dataset 'Employer\_survey\_dataset\_July24' to create the final weighted sample dataset 'VNMWtd-Employer\_survey\_dataset\_Final with PopnWtFin'.

Table 3 provides a brief description of the new variables that were added to the final weighted sample dataset 'VNMWtd-Employer\_survey\_dataset\_Final with PopnWtFin'.

<b>Table 3: Description of variables added to the Viet Nam Sample Dataset 'VNMWtd-Employer_survey_dataset_Final with PopnWtFin'</b>	
<b>Variable Name</b>	<b>Description</b>
Sample	Sample Type Indicator for Initial sample and Reserve Sample
Code	Response Code; 1-Complete, 2-Ineligible, 3-Refusal, 4-Other Nonresponse
Note	Response Alphabetic Description; Complete, Ineligible, Refusal, Other Nonresponse
Size	Firm Size Stratum; 1-1 to10 employees, 2-11 to 50 employees, 3-51 or more employees
PopnStr	Stratum Population Count
SmplStr	Stratum Final Sample Count
TheorWtStr	Stratum Theoretical weight based on the sample design
nComplStr	Stratum count of number of complete cases in the sample
IneligStrCnt	Stratum count of number of ineligible cases in the sample
NoRespStrCnt	Stratum count of number of non-response cases in the sample
PopnStrAdj	Stratum population size that has been adjusted for the number of ineligible cases in the sample
PopnWtFin	Stratum Final Population Weight resulting from the adjustment of the stratum Theoretical weight for the number of ineligible cases in the sample and the number of non-response cases in the sample.  <b>This is the sample weight to be used to produce survey estimates.</b>

## 4 Construction of the Final Sample Weights

### 4.1.1 Overview

The Viet Nam Employer Survey data was weighted in order to correct for imperfections in the sample that might lead to bias between the sample and the target population. Such imperfections include the selection of firms with unequal probabilities between strata, sample frame imperfections such as ineligible and duplicate firm listings, and non-response. In other words, the purposes of weighting are:

- 1) To compensate for unequal probabilities of selection between strata;
- 2) To adjust the target population size for ineligible and duplicate firm listings;
- 3) To compensate for non-responding firms.

The weights were constructed as follows:

- 1) A Theoretical Design Weight, *TheorWtStr*, was calculated for each of the 594 sampled firms. This is a stratum weight based on the stratified sample design. Each sampled firm's Theoretical Design weight is the number of firms that each sampled firm represents in the sample frame. This weight is the inverse of the probability of selection in a stratum.
- 2) The sample frame stratum population counts, *PopnStr*, were adjusted to take into account the number of ineligible firms in each stratum sample. The resulting adjusted stratum population size is the variable '*PopnStrAdj*'.
- 3) A final population stratum weight, *PopnWtFin*, was constructed by adjusting the Theoretical Design Weight, *TheorWtStr* for
  - a. Number of ineligible firms in the sample
  - b. Number of non-responding firms in the sample

### 4.1.2 Theoretical Design Weight

The development of the sampling weights starts with the construction of the base weight, or theoretical design weight, for each sampled unit to correct for the unequal probabilities of selection between the three Firm Size strata. In general, the base weight of a sampled firm is the inverse of its probability of selection into the sample.

In mathematical notation, the basic theoretical design weight, *TheorWtStr*, for each sampled firm 'i' was calculated as,

$$TheorWtStr = \frac{N_h}{n_h} \dots \dots \dots [1]$$

where,

$N_h$  is the number of firms in stratum 'h' in the sample frame,  
 $n_h$  is the number of sampled firms in stratum 'h'.

Since the sampled firms were selected with equal probability within a stratum the theoretical design weight, *TheorWtStr*, is the same for each sampled firm within a stratum.

Table 4 shows the theoretical design weight, *TheorWtStr*, for each stratum of the Viet Nam Employer Survey.

Note: The theoretical design weight, *TheorWtStr*, should not be used for producing survey estimates. These weights were calculated as the basis for creating the final population stratum weights.

<b>Table 4: TheorWtStr by Stratum</b>			
<b>Stratum – Firm Size (# of employees)</b>	<b>Population Size</b>	<b>Final Sample Size</b>	<b>TheorWtStr (approx)</b>
<b>1 to 10</b>	16,519	189	87.40
<b>11 to 50</b>	8,631	200	43.16
<b>51 or more</b>	4,764	205	23.24
<b>Total</b>	<b>29,914</b>	<b>594</b>	<b>n/a</b>

#### 4.1.3 Sample Frame Population Adjustment for Ineligible Firms

The 594 firms selected from the sample frame 'Firms in scope' included a number of firms that were ineligible for the Viet Nam Employer Survey. Consequently, the sample frame stratum population counts, *PopnStr*, were proportionally adjusted to take into account the number of ineligible firms that were included in the sample. The resulting adjusted stratum population count, *PopnStrAdj*, is the estimated actual number of eligible firms in the sample frame from which the sample was selected.

The adjustment to the sample frame stratum population counts to take into account the number of sampled ineligible firms was calculated as,

$$PopStrAdj = N_h \times \left(1 - \frac{n_{ex,h}}{n_h}\right) \dots \dots \dots [2]$$

where,

$N_h$  is the number of firms in stratum 'h' in the sample frame,

$n_{ex,h}$  is the number of sampled ineligible firms in stratum 'h',

$n_h$  is the number of sampled firms in stratum 'h'.

Table 5 provides the population stratum counts adjusted for the number of sampled ineligible firms.

<b>Table 5: Estimated actual number of eligible firms PopnStrAdj by Stratum</b>			
<b>Stratum – Firm Size (# of employees)</b>	<b>Original Population Size</b>	<b>Final Sample Size</b>	<b>PopnStrAdj Population count adjusted for ineligible firms</b>
<b>1 to 10</b>	16,519	189	12,936
<b>11 to 50</b>	8,631	200	7,552
<b>51 or more</b>	4,764	205	4,508
<b>Total</b>	<b>29,914</b>	<b>594</b>	<b>24,996</b>

#### 4.1.4 Final Stratum Population Weight – Adjustment for Non-responding Firms

The weighting of the Viet Nam Employer Survey concludes with the calculation of the final stratum population weight, **PopnWtFin**, which is the weight to be used to produce the population estimates of the characteristics of interest from the sample dataset.

The final stratum population weight, **PopnWtFin**, for each sampled firm takes into account the number of non-responding sampled firms as well as the number of ineligible sampled firms. **PopnWtFin** was calculated as,

$$PopnWtFin = \begin{cases} \frac{PopStrAdj}{n_{c,h}}, & Code = 1 \\ 0, & Code \neq 1 \end{cases} \dots\dots\dots [3]$$

where,

$n_{c,h}$  is the number of complete interviews from the sampled firms in stratum 'h'.

## 5 Response Rate

The calculation of the response rate for the Viet Nam Employer Survey uses the following terms derived from the sample of 594 firms:

$n_h$  = number of sampled firms from stratum  $h$

$n_{ex,h}$  = number of sampled firms that should have been excluded from stratum  $h$

$n_{c,h}$  = number of firms in stratum  $h$  that participated

$n_{nr,h}$  = number of non-responding firms in stratum  $h$

Note that the following equation holds:

$$n_{ex,h} + n_{c,h} + n_{nr,h} = n_h$$

The stratum response rate  $R_h$  is defined as the ratio of sampled firms that participated to the total number of sampled firms minus any excluded firms. It was calculated by the following equation:

$$R_h = \frac{n_{c,h}}{n_{c,h} + n_{nr,h}} \dots \dots \dots [4]$$

Table 6 provides the response rates by Firm Size, as well as the overall response rate.

<b>Stratum – Firm Size (# of employees)</b>	<b>Complete</b>	<b>Non-response</b>	<b>Total Eligible</b>	<b>Response Rate</b>
(1)	(2)	(3)	(4)=(2)+(3)	(2)/(4)
<b>1 to 10</b>	87	61	148	58.8%
<b>11 to 50</b>	114	61	175	65.1%
<b>51 or more</b>	129	65	194	66.5%
<b>Total</b>	<b>330</b>	<b>187</b>	<b>517</b>	<b>63.8%</b>