The Vietnam 2015 Enterprise Surveys Data Set

I. Introduction

This document provides additional information on the data collected in Vietnam between November 2014 and April 2016. The objective of the Enterprise Survey is to gain an understanding of what firms experience in the private sector.

As part of its strategic goal of building a climate for investment, job creation, and sustainable growth, the World Bank has promoted improving the business environment as a key strategy for development, which has led to a systematic effort in collecting enterprise data across countries. The Enterprise Surveys (ES) are an ongoing World Bank project in collecting both objective data based on firms' experiences and enterprises' perception of the environment in which they operate.

The ES currently cover over 130,000 firms in 135 countries, of which 121 have been surveyed following the standard methodology. This allows for better comparisons across countries and across time. Data are used to create statistically significant business environment indicators that are comparable across countries. The ES are also used to build a panel of enterprise data that will make it possible to track changes in the business environment over time and allow, for example, impact assessments of reforms.

This report outlines and describes the sampling design of the data, the data set structure as well as additional information that may be useful when using the data, such as information on non-response cases and the appropriate use of the weights.

II. Sampling Structure

The sample for 2015 Vietnam ES was selected using stratified random sampling, following the methodology explained in the *Sampling Note*[1]. Stratified random sampling[2] was preferred over simple random sampling for several reasons[3]:

a. To obtain unbiased estimates for different subdivisions of the population with some known level of precision.

b. To obtain unbiased estimates for the whole population. The whole population, or universe of the study, is the non-agricultural economy. It comprises: all manufacturing sectors according to the group classification of ISIC Revision 3.1: (group D), construction sector (group F), services sector (groups G and H), and transport, storage, and communications sector (group I). Note that this definition excludes the following sectors: financial intermediation (group J), real estate and renting activities (group K, except sub-sector 72, IT, which was added to the population under study), and all public or utilities-sectors.

---

[1] The complete text can be found at
http://www.enterprisesurveys.org/~/media/GIAWB/EnterpriseSurveys/Documents/Methodology/Sampling_Note.pdf

[2] A stratified random sample is one obtained by separating the population elements into non-overlapping groups, called strata, and then selecting a simple random sample from each stratum. (Richard L. Scheaffer; Mendenhall, W.; Lyman, R., "Elementary Survey Sampling", Fifth Edition).

[3] Cochran, W., 1977, pp. 89; Lohr, Sharon, 1999, pp. 95

c. To make sure that the final total sample includes establishments from all different sectors and that it is not concentrated in one or two of industries/sizes/regions.

d. To exploit the benefits of stratified sampling where population estimates, in most cases, will be more precise than using a simple random sampling method (i.e., lower standard errors, other things being equal.)

e. Stratification may produce a smaller bound on the error of estimation than would be produced by a simple random sample of the same size. This result is particularly true if measurements within strata are homogeneous.

f. The cost per observation in the survey may be reduced by stratification of the population elements into convenient groupings.

Three levels of stratification were used in this country: industry, establishment size, and region. The original sample design with specific information of the industries and regions chosen is described in Appendix C.

Industry stratification was designed in the way that follows: the universe was stratified into five manufacturing industries and two services industries- Food and Beverages (ISIC Rev. 3.1 code 15), Garments (ISIC code 18), Non-metallic mineral products (ISIC code 26), Fabricated metal products (ISIC code 28), Other Manufacturing (ISIC codes 16,17, 19-25, 27, 29-37), Retail (ISIC code 52) and Other Services (ISIC codes 45, 50, 51, 55, 60-64, and 72).

For the Vietnam ES, size stratification was defined as follows: small (5 to 19 employees), medium (20 to 99 employees), and large (100 or more employees).

Regional stratification for the Vietnam ES was done across four regions: Red River Delta, North Central Area and Central Coastal Area, South East, and Mekong River Delta.

III. Sampling implementation
Given the stratified design, sample frames containing a complete and updated list of establishments as well as information on all stratification variables (number of employees, industry, and region) are required to draw the sample. Great efforts were made to obtain the best source for these listings.

Mekong Economics was the contractor that implemented the Vietnam 2015 ES.

The sample frame consisted of listings of firms from two sources: First, for panel firms the list of 1053 firms from the Vietnam 2009 ES was used and second, for fresh firms (i.e., firms not covered in 2009), data from the General Statistics Office (GSO) of Vietnam was used.

Table 1: Vietnam ES Sample Frame (Fresh and Panel Combined)

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	56	145	74	93	174	91	148	**2,045**
	Medium	50	78	63	59	301	41	144	
	Large	53	66	80	129	92	51	57	
North Central area and Central coastal area	Small	48	40	148	60	72	50	117	**1,176**
	Medium	52	48	64	40	43	48	62	
	Large	53	53	57	22	35	25	39	
South East	Small	215	220	178	135	261	379	337	**3,532**
	Medium	128	117	147	75	207	49	243	
	Large	116	257	101	72	158	89	48	
Mekong River Delta	Small	137	21	33	66	67	44	90	**1,142**
	Medium	122	45	37	46	48	45	37	
	Large	145	59	18	20	22	14	26	
		1,175	**1,149**	**1,000**	**817**	**1,480**	**926**	**1,348**	**7,895**

Source: World Bank and General Statistics Office of Vietnam

Table 2: Vietnam Sample Frame (Panel)

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	6	2	4	5	21	17	20	**333**
	Medium	14	10	13	14	52	11	22	
	Large	11	21	23	12	46	1	8	
North Central area and Central coastal area	Small	4	0	2	8	10	14	18	**226**
	Medium	11	3	12	7	28	10	24	
	Large	11	10	14	1	20	1	18	
South East	Small	10	7	5	6	18	13	19	**378**
	Medium	15	26	18	9	52	5	17	
	Large	16	32	20	12	69	3	6	
Mekong River Delta	Small	12	0	2	8	1	7	11	**116**
	Medium	5	1	2	8	11	6	10	
	Large	8	6	0	4	7	1	6	
		123	**118**	**115**	**94**	**335**	**89**	**179**	**1,053**

The quality of the frame was enhanced by the verification process conducted by Mekong Economics. However, the sample frame was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc.

Given the impact that non-eligible units included in the sample universe may have on the results, adjustments may be needed when computing the appropriate weights for individual observations. The percentage of confirmed non-eligible units as a proportion of the total number of sampled establishments contacted for the survey was 7.3% (285 out of 3889 establishments)[4].

Breaking down by industry and size, the following sample targets were achieved (based on the sampling information):

---

[4] Based on out of target and ineligible contacts

Table 3: Achieved Interviews (Fresh and Panel Combined)

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Other services
Red River Delta	Small	12	12	14	20	23	19	22	**302**
	Medium	13	12	16	15	15	10	20	
	Large	13	14	17	12	7	3	13	
North Central area and Central coastal area	Small	13	12	7	13	6	14	22	**240**
	Medium	11	12	16	11	13	9	12	
	Large	12	12	15	10	7	7	6	
South East	Small	9	14	8	17	29	19	24	**307**
	Medium	13	19	16	10	22	5	21	
	Large	11	19	17	12	10	4	8	
Mekong River Delta	Small	16	3	8	11	3	9	7	**147**
	Medium	12	12	4	7	6	6	6	
	Large	8	8	3	5	5	3	5	
		143	**149**	**141**	**143**	**146**	**108**	**166**	**996**

Table 4: Achieved Interviews (Panel)

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	2	2	2	3	8	6	2	**93**
	Medium	5	5	6	7	7	4	2	
	Large	6	7	7	5	4	1	2	
North Central area and Central coastal area	Small	1	0	0	2	3	5	2	**68**
	Medium	3	2	5	4	10	5	2	
	Large	5	6	6	1	4	0	2	
South East	Small	1	2	0	2	12	6	2	**105**
	Medium	7	11	7	4	8	2	4	
	Large	5	10	8	6	4	0	4	
Mekong River Delta	Small	0	0	2	3	0	2	2	**28**
	Medium	1	0	0	1	3	3	2	
	Large	2	1	0	1	3	0	2	
		38	**46**	**43**	**39**	**66**	**34**	**28**	**294**

IV. Data Base Structure:

The structure of the data base reflects the fact that 2 different versions of the survey instrument were used for all registered establishments. Questionnaires have common questions (*core* module) and respectfully additional manufacturing- and services-specific questions. The eligible manufacturing industries have been surveyed using the ***Manufacturing*** questionnaire (includes the *core* module, plus manufacturing specific questions). Retail firms have been interviewed using the ***Services*** questionnaire (includes the *core* module plus retail specific questions) and the residual eligible services have been covered using the ***Services*** questionnaire (includes the *core* module). Each variation of the questionnaire is identified by the index variable, *a0*.

All variables are named using, first, the letter of each section and, second, the number of the variable within the section, i.e. *a1* denotes section *A*, question *1* (some exceptions apply due to comparability reasons). Variable names preceded by the prefix *"EA"* or "MYA" indicate questions specific to Vietnam and other countries in EAP 2015, therefore, they may not be found in the implementation of the rollout in other countries. All other suffixed variables are global and are present in all country surveys over the world. All variables are numeric with the exception of those variables with an "x" at the end of their names. The suffix "x" denotes that the variable is alpha-numeric.

There are 2 establishment identifiers, *idstd* and *id*. The first is a global unique identifier. The second is a country unique identifier. The variables *a2* (sampling region), *a6a* (sampling establishment's size), and *a4a* (sampling sector) contain the establishment's classification into the strata chosen for each country using information from the sample frame. The strata were defined according to the guidelines described above.

There are three levels of stratification: industry, size and region. Different combinations of these variables generate the strata cells for each industry/region/size combination. A distinction should be made between the variable a4a and d1a2 (industry expressed as ISIC rev. 3.1 code). The former gives the establishment's classification into one of the chosen industry-strata based on the sample frame, whereas the latter gives the establishment's actual industry classification (four digit code) based on the main activity at the time of the survey.

All of the following variables contain information from the sampling frame. They may not coincide with the reality of individual establishments as sample frames may contain inaccurate or outdated information. The variables containing the sample frame information are included in the data set for researchers who may want to further investigate statistical features of the survey and the effect of the survey design on their results.

-*a2* is the variable describing sampling regions

-*a6a*: coded using the same standard for small, medium, and large establishments as defined above.

-*a4a*: coded following the stratification by sector as defined above.

The surveys were implemented following a 2 stage procedure. Typically first a screener questionnaire is applied over the phone to determine eligibility and to make appointments. Then a face-to-face interview takes place with the Manager/Owner/Director of each establishment. However, sometimes the phone numbers were unavailable in the

sample frame, and thus the enumerators applied the screeners in person. The variables *a4b* and *a6b* contain the industry and size of the establishment from the screener questionnaire. Variables *a8* to *a11* contain additional information and were also collected in the screening phase.

Note that there are variables for size (*l1*, *l6* and *l8*) that reflect more accurately the reality of each establishment. Advanced users are advised to use these variables for analytical purposes. Variables *l1* (number of permanent full-time workers at the end of the last complete fiscal year), *l6* (number of full-time seasonal workers employed during last complete fiscal year) and *l8* (average length of employment of full-time temporary employees during last complete fiscal year) were designed to obtain a more accurate measure of employment accounting for permanent and temporary employment. Special efforts were made to make sure that this information was not missing for most establishments.

Most of the firms interviewed used January 2014 to December 2014 as their fiscal year. Variables *eaa3a3y* (last complete fiscal year) and *eaa3a3w* (starting month of last complete fiscal year) can be used to obtain the last complete fiscal year for each firm.

A small subset of questions denoted by ""MYA" were asked to 370 firms as a paper-based phone follow-up instead of being part of the main interview. The variable *eava0* can be used to track these. It takes the value of 1 when the follow-up was administered successfully, and 2 when the respondent refused to answer these questions. For the rest of the respondents this portion of the interview was conducted as part of the main interview, and the variable eava0 has no entries.

For questions pertaining to monetary amounts, the unit is the Vietnamese dong.

V. Universe Estimates

Universe estimates for the number of establishments in each cell in Vietnam were produced for the strict, weak and median eligibility definitions described below. The estimates were the multiple of the relative eligible proportions.

For some establishments where contact was not successfully completed during the screening process (because the firm has moved and it is not possible to locate the new location, for example), it is not possible to directly determine eligibility. Thus, different assumptions about the eligibility of establishments result in different adjustments to the universe cells and thus different sampling weights.

Three sets of assumptions on establishment eligibility are used to construct sample adjustments using the status code information.

Strict assumption: eligible establishments are only those for which it was possible to directly determine eligibility. The resulting weights are included in the variable *wstrict*.

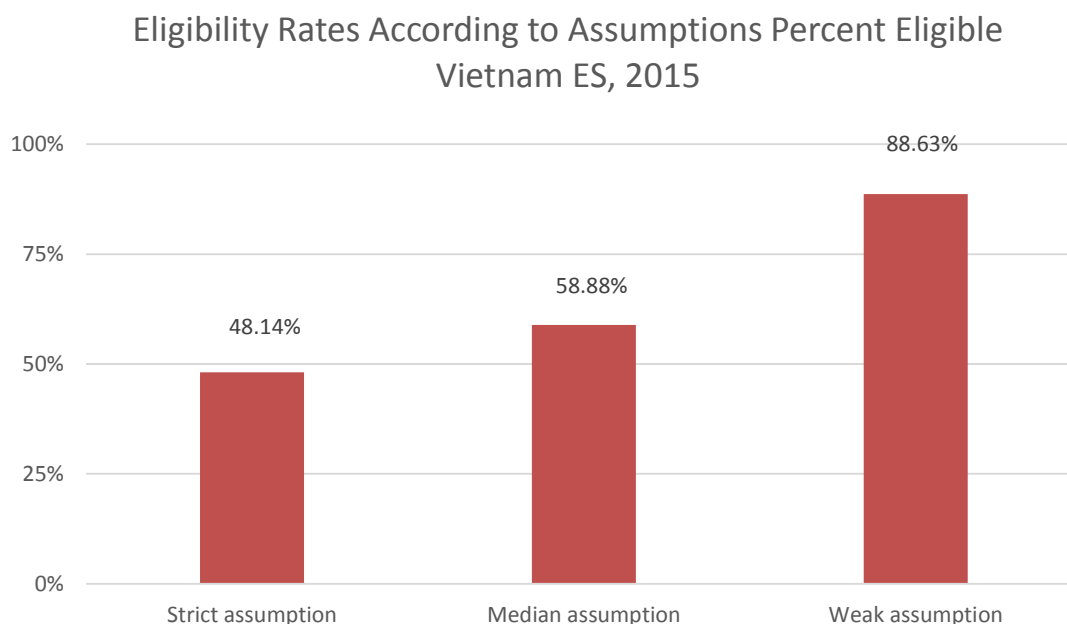*Strict eligibility = (Sum of the firms with codes 1,2,3,4,&16) / Total*

Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

Median eligibility = (Sum of the firms with codes 1,2,3,4,16,10,11, & 13) / Total

Weak assumption: in addition to the establishments included in points a and b, all establishments for which it was not possible to contact or that refused the screening questionnaire are assumed eligible. This definition includes as eligible establishments with dead or out of service phone lines, establishments that never answered the phone, and establishments with incorrect addresses for which it was impossible to find a new address. Under the weak assumption only observed non-eligible units are excluded from universe projections. The resulting weights are included in the variable *wweak*.

Weak eligibility= (Sum of the firms with codes, 1,2,3,4,16,10,11,13,91,92,93,94,12) / Total

The indicators computed for the ES website use the median weights. The following graph shows the different eligibility rates calculated for firms in the sample frame under each set of assumptions.

Eligibility Rates According to Assumptions Percent Eligible Vietnam ES, 2015

Universe estimates for the number of establishments in each industry-region-size cell in Vietnam were produced for the strict, weak and median eligibility definitions. Appendix B shows the universe estimates of the numbers of registered establishments that fit the criteria of the ES.

Once an accurate estimate of the universe cell projection was made, weights for the probability of selection were computed using the number of completed interviews for each cell.

VI. Weights

Since the sampling design was stratified and employed differential sampling, individual observations should be properly weighted when making inferences about the population. Under stratified random sampling, unweighted estimates are biased unless sample sizes are proportional to the size of each stratum. With stratification the probability of selection of each unit is, in general, not the same. Consequently, individual observations must be weighted by the inverse of their probability of selection (probability weights or *pw* in Stata.)[5]

Special care was given to the correct computation of the weights. It was imperative to accurately adjust the totals within each region/industry/size stratum to account for the presence of ineligible units (the firm discontinued businesses or was unattainable, education or government establishments, no reply after having called in different days of the week and in different business hours, no tone in the phone line, answering machine, fax line[6], wrong address or moved away and could not get the new references). The information required for the adjustment was collected in the first stage of the implementation: the screening process. Using this information, each stratum cell of the universe was scaled down by the observed proportion of ineligible units within the cell. Once an accurate estimate of the universe cell (projections) was available, weights were computed using the number of completed interviews.

VII. Appropriate use of the weights

Under stratified random sampling, weights should be used when making inferences about the population. Any estimate or indicator that aims at describing some feature of the population should take into account that individual observations may not represent equal shares of the population.

However, there is some discussion as to the use of weights in regressions (see Deaton, 1997, pp.67; Lohr, 1999, chapter 11, Cochran, 1953, pp.150). There is not strong large-sample econometric argument in favor of using weighted estimation for a common population coefficient if the underlying model varies per stratum (stratum-specific coefficient): both simple OLS and weighted OLS are inconsistent under regular conditions. However, weighted OLS have the advantage of providing an estimate that is independent of the sample design. This latter point may be quite relevant for the ES as in most cases the objective is not only to obtain model-unbiased estimates but also design-unbiased estimates (see also Cochran, 1977, pp 200 who favors the used of weighted OLS for a common population coefficient.)[7]

---

[5] This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

[6] For the surveys that implemented a screener over the phone.

[7] Note that weighted OLS in Stata using the command regress with the option of weights will estimate wrong standard errors. Using the Stata survey specific commands *svy* will provide appropriate standard errors.
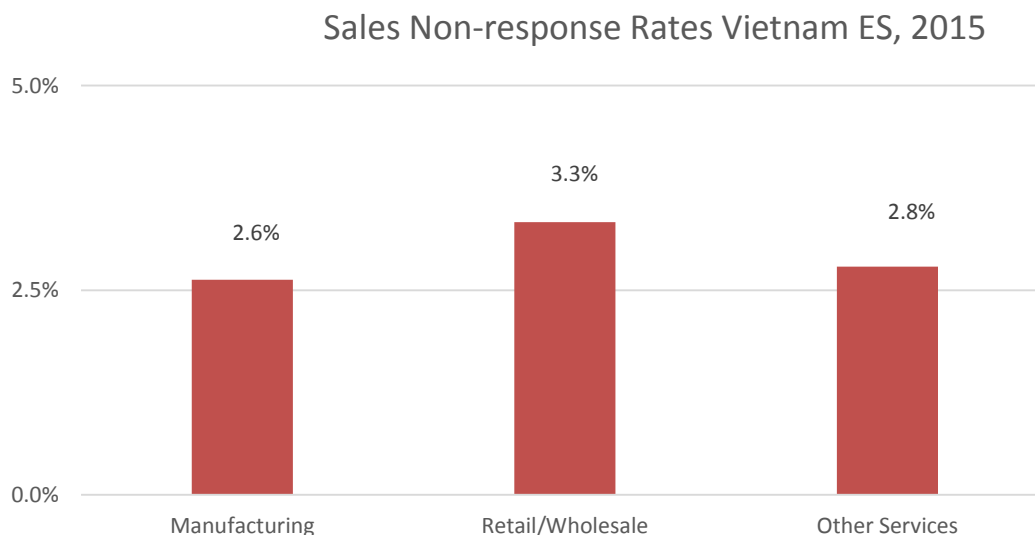
From a more general approach, if the regressions are descriptive of the population then weights should be used. The estimated model can be thought of as the relationship that would be expected if the whole population were observed.[8] If the models are developed as structural relationships or behavioral models that may vary for different parts of the population, then, there is no reason to use weights.

VIII. Non-response

Survey non-response must be differentiated from item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the refusals to answer some specific questions. Enterprise Surveys suffer from both problems and different strategies were used to address these issues.

Item non-response was addressed by two strategies:
a- For sensitive questions that may generate negative reactions from the respondent, such as corruption or tax evasion, enumerators were instructed to collect the refusal to respond (-8) as a different option from don't know (-9).
b- Establishments with incomplete information were re-contacted in order to complete this information, whenever necessary. However, there were clear cases of low response. The following graph shows non-response rates for the sales variable, *d2*, by sector. Please, note that for this specific question, refusals were not separately identified from "Don't know" responses.

Sales Non-response Rates Vietnam ES, 2015

Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) was suggested for interview. Survey non-response did occur but substitutions were made in order to potentially achieve strata-specific goals; whenever this was done, strict rules were followed to ensure replacements

---

[8] The use weights in most model-assisted estimations using survey data is strongly recommended by the statisticians specialized on survey methodology of the JPSM of the University of Michigan and the University of Maryland.

were randomly selected within the same stratum. Further research is needed on survey non-response in the Enterprise Surveys regarding potential introduction of bias.

As the following graph shows, the number of interviews per contacted establishments was 0.26.[9] This number is the result of two factors: explicit refusals to participate in the survey, as reflected by the rate of rejection (which includes rejections of the screener and the main survey) and the quality of the sample frame, as represented by the presence of ineligible units. The share of rejections per contact was 0.31.

Rejection rate and Interviews per Contact
Vietnam ES, 2015

Details on the rejection rate, eligibility rate, and item non-response are available at the level strata. This report summarizes these numbers to alert researchers of these issues when using the data and when making inferences. Item non-response, selection bias, and faulty sampling frames are not unique to Vietnam. All enterprise surveys suffer from these shortcomings, but in very few cases they have been made explicit.

References:

Cochran, William G., Sampling Techniques, New York, New York: John Wiley & Sons, 1977.

Deaton, Angus, The Analysis of Household Surveys, Baltimore, Maryland: Johns Hopkins University Press, 1998.

Levy, Paul S. and Stanley Lemeshow, Sampling of Populations: Methods and Applications, New York, New York: John Wiley & Sons, 1999.

Lohr, Sharon L. Samping: Design and Techniques, Boston, Massachusetts: Brookes/Cole, 1999.

Scheaffer, Richard L.; Mendenhall, W.; Lyman, R., Elementary Survey Sampling, Fifth Edition, 1996.

---

[9] The estimate is based on the total no. of firms contacted including ineligible establishments.

Appendix A

Status Codes Enterprise Survey (ES):

157	Screening in process	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	157

1872	Eligible	1. Eligible establishment (Correct name and address)	1815
		2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	8
		3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	12
		4. Eligible establishment (Moved and traced)	34
		16. Eligible establishment (Panel Firm - now less than five employees; this code applies only to panel firms.)	3

333	Screener refusal	13. Refuses to answer the screener	333

146	Ineligible	5. The establishment has less than 5 permanent full time employees	4
		616. The firm discontinued businesses - (Establishment went bankrupt)	79
		617.	0
		618. The firm discontinued businesses - (Original establishment disappeared and is now a different firm)	4
		619. The firm discontinued businesses - (Establishment was bought out by another firm)	4
		620. The firm discontinued businesses - (It was impossible to determine for what reason)	47
		621. The firm discontinued businesses - (Other)	5
		7. Not a business: Private household	0
		8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	3
139	Out of target	151. Out of target - outside the covered regions	3
		152. Out of target - moved abroad	0
		153. Out of target - Not registered with Statistical Authority	122

		154. Out of target - establishment is HQ without production or sales of goods or services	0
		155. Out of target - establishment was not in operation for the entirety of last fiscal year	0
		156. Duplicated firm within the sample	14
1242	**Unobtainable**	91. No reply after having called in different days of the week and in different business hours	227
		92. Line out of order	123
		93. No tone	14
		94. Phone number does not exist	372
		10. Answering machine	33
		11. Fax line- data line	52
		12. Wrong address/ moved away and could not get the new references	421

3889	**Total contacted**

Response Outcomes: Vietnam ES 2015:

Target and totals	Sample target	1000
	Sample target completion rate	99.6%
	Total contacts available in frame	7895
	Total contacts issued	7895
	Total contacts contacted	3889

Screening phase	Screening in process	157
	Eligibles	1872

15

	Screener refusal	333
	Ineligible + out of target	285
	Unobtainable	1242
Interview phase (only if eligible)	Complete interviews without extra module	996
	Complete interviews with extra module	0
	Eligible in process + incomplete interviews	4
	Interview refusal	861

	Screening in process rate	4.0%
	Screener refusal rate	8.6%
Percent breakdown (relative to total contacted)	Ineligible + out of target rate	7.3%
	Unobtainable rate	31.9%
	Interview conversion rate	25.6%
	Eligible in process + incomplete interviews rate	0.1%
	Interview refusal rate	22.1%

Appendix B: Universe Estimate Based on Sampling Weights, Vietnam ES 2015

Strict Universe Estimates – Fresh:

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	319	147	126	780	1,620	1,716	15,352	**30,923**
	Medium	163	209	185	389	1,161	275	6,012	
	Large	86	213	125	81	545	33	1,386	
North Central area and Central coastal area	Small	195	54	86	244	506	1,456	5,816	**13,193**
	Medium	101	63	131	88	322	189	2,854	
	Large	94	60	67	17	174	23	655	
South East	Small	287	304	86	547	1,774	1,514	12,255	**25,608**
	Medium	249	301	116	305	1,419	212	3,906	
	Large	177	313	52	103	856	44	788	
Mekong River Delta	Small	402	18	37	128	294	1,014	2,665	**6,874**
	Medium	321	34	41	36	230	101	1,023	
	Large	186	43	17	10	116	12	147	
		2,579	**1,761**	**1,068**	**2,728**	**9,016**	**6,587**	**52,859**	**76,598**

Median Universe Estimates – Fresh:

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	411	181	162	1,093	2,140	2,159	18,265	**37,334**
	Medium	201	247	227	523	1,474	332	6,870	
	Large	110	261	160	113	719	41	1,645	
North Central area and Central coastal area	Small	239	64	105	326	637	1,747	6,596	**15,161**

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
	Medium	119	71	153	113	390	218	3,109	
	Large	115	70	82	23	219	27	741	
South East	Small	370	374	110	766	2,347	1,907	14,595	**31,159**
	Medium	308	356	143	411	1,803	256	4,468	
	Large	227	384	66	144	1,130	55	937	
Mekong River Delta	Small	516	22	47	179	387	1,272	3,163	**8,366**
	Medium	395	40	50	48	291	122	1,166	
	Large	238	53	21	15	152	15	174	
		3,248	**2,123**	**1,326**	**3,753**	**11,689**	**8,152**	**61,728**	**92,020**

Weak Universe Estimates – Fresh:

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	612	278	321	1,623	3,683	3,467	28,859	**56,424**
	Medium	259	329	391	673	2,196	461	9,398	
	Large	128	315	248	132	967	51	2,033	
North Central area and Central coastal area	Small	391	108	228	531	1,204	3,081	11,453	**25,137**
	Medium	169	104	290	159	638	333	4,674	
	Large	147	92	139	29	323	38	1,005	
South East	Small	642	671	254	1,325	4,701	3,564	26,846	**54,934**
	Medium	463	552	287	615	3,127	415	7,116	
	Large	308	538	120	195	1,770	80	1,347	
Mekong River Delta	Small	796	35	97	275	690	2,115	5,174	**13,054**
	Medium	529	56	88	64	449	176	1,651	
	Large	287	66	34	18	212	20	223	
		4,731	**3,144**	**2,497**	**5,637**	**19,959**	**13,801**	**99,780**	**149,549**

		Food and Beverages	Garments	Non metallic mineral products	Fabricated metal products	Other manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	612	278	321	1,623	3,683	3,467	28,859	**56,424**
	Medium	259	329	391	673	2,196	461	9,398	
	Large	128	315	248	132	967	51	2,033	
North Central area and Central coastal area	Small	391	108	228	531	1,204	3,081	11,453	**25,137**
	Medium	169	104	290	159	638	333	4,674	
	Large	147	92	139	29	323	38	1,005	
South East	Small	642	671	254	1,325	4,701	3,564	26,846	**54,934**
	Medium	463	552	287	615	3,127	415	7,116	
	Large	308	538	120	195	1,770	80	1,347	
Mekong River Delta	Small	796	35	97	275	690	2,115	5,174	**13,054**
	Medium	529	56	88	64	449	176	1,651	
	Large	287	66	34	18	212	20	223	
		4,731	**3,144**	**2,497**	**5,637**	**19,959**	**13,801**	**99,780**	**149,549**

Appendix C: Original Sample Design, Vietnam ES 2015

Original Sample Design (Fresh)

		Food	Garments	Non metallic Mineral Products	Fabricated Metal Products	Other Manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	9	11	10	16	13	13	17	**189**
	Medium	6	7	7	8	8	6	13	
	Large	6	7	7	6	5	7	7	
North Central area and Central coastal area	Small	9	12	11	6	4	9	8	**140**
	Medium	6	9	7	5	3	4	7	
	Large	6	6	7	9	3	6	3	
South East	Small	7	9	9	13	18	13	17	**188**
	Medium	7	7	6	7	12	5	11	
	Large	7	7	7	6	9	6	5	
Mekong River Delta	Small	7	13	11	6	4	7	4	**137**
	Medium	9	11	11	6	2	4	3	
	Large	7	6	12	6	2	4	2	
		86	**105**	**105**	**94**	**83**	**84**	**97**	**654**

Original Sample Design (Panel)

		Food	Garments	Non metallic Mineral Products	Fabricated Metal Products	Other Manufacturing	Retail	Other Services	Grand Total
Red River Delta	Small	6	2	4	5	3	4	2	**94**
	Medium	6	7	7	7	3	4	2	
	Large	6	7	7	6	3	1	2	

North Central area and Central coastal area	Small	4	0	2	7	3	4	2	**78**
	Medium	6	3	7	6	3	4	2	
	Large	6	6	6	1	3	1	2	
South East	Small	8	7	5	6	4	4	2	**110**
	Medium	7	8	7	7	4	4	2	
	Large	6	8	6	6	4	3	2	
Mekong River Delta	Small	8	0	2	6	1	3	2	**64**
	Medium	5	1	2	5	3	3	2	
	Large	6	6	0	4	3	1	1	
		74	**55**	**55**	**66**	**37**	**36**	**23**	**346**