

# **The Socialist Republic of Vietnam 2009 Enterprise Surveys Data Set**

## **1. Introduction**

1. This document provides additional information on the data collected in Vietnam between June 2009 and January 2010 as part of the Enterprise Survey component of the survey East Asia and Pacific Enterprise Survey 2009.

The objective of the survey is to obtain feedback from enterprises on the state of the private sector in client countries (Lao PDR, Tonga, Samoa, Vanuatu, Timor Leste, Fiji, Federated States of Micronesia, Philippines, Vietnam, Indonesia, and Papua New Guinea) as well as to help in building a panel of enterprise data that will make it possible to track changes in the business environment over time, thus allowing, for example, impact assessments of reforms.

Through interviews with firms in the manufacturing and services sectors, the survey will assess the constraints to private sector growth and create statistically significant business environment indicators that are comparable across countries.

The 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.

## **2. Sampling Structure**

2. The sample for Vietnam was selected using stratified random sampling, following the methodology explained in the Sampling Manual<sup>1</sup>. Stratified random sampling<sup>2</sup> was preferred over simple random sampling for several reasons<sup>3</sup>:

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.

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.)

---

<sup>1</sup> The complete text can be found at [http://www.enterprisesurveys.org/documents/Implementation\\_note.pdf](http://www.enterprisesurveys.org/documents/Implementation_note.pdf)

<sup>2</sup> 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).

<sup>3</sup> Cochran, W., 1977, pp. 89; Lohr, Sharon, 1999, pp. 95

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.

3. 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 E.

4. Industry stratification was designed in the way that follows: the universe was stratified into 5 manufacturing industries, 1 services industry -retail -, and two residual sectors as defined in the sampling manual. Each manufacturing industry had a target of 120 - 145 interviews. The services industry and the two residual sectors had a target of 120 interviews. For the manufacturing industries sample sizes were inflated by about 25% to account for potential non-response cases when requesting sensitive financial data and also because of likely attrition in future surveys that would affect the construction of a panel.

5. Size stratification was defined following the standardized definition for the rollout: small (5 to 19 employees), medium (20 to 99 employees), and large (more than 99 employees). For stratification purposes, the number of employees was defined on the basis of reported permanent full-time workers. This seems to be an appropriate definition of the labor force since seasonal/casual/part-time employment is not a common practice, except in the sectors of construction and agriculture.

6. Regional stratification was defined in five regions containing 14 provinces: Red River Delta (Hanoi, Ha Tay, Hai Duong, and Hai Phong), the North Centre Coast (Thanh Hoa, Nghe An), Mekong River Delta (Can Tho, Long An, Tien Giang), South Centre Coast (Khanh Hoa, Da Nang) and South East (Ho Chi Minh City, Binh Duong, Dong Nai).

### **3. Sampling implementation**

7. 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 for the Indicator Surveys.

8. Two frames were used for Vietnam. The sample frame containing fresh contacts used in the Vietnam was obtained from the 2008 Vietnam General Statistics Office. A frame containing firms that had participated in the 2005 survey constituted a second frame of panel contacts. Each database contained the following information:

- Name of the firm
- Location
- Contact details
- ISIC code

-Number of employees.

Counts from sample frames shown below.

### Universe Figures for Vietnam

ELIGIBLE UNIVERSE (Fresh)										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	46	3	8	35	41	301	129	1483	2046
	2-Medium	25	5	6	116	15	35	122	768	1092
	3-Large	23	3	11	54	6	4	39	155	295
	TOTAL	94	11	25	205	62	340	290	2406	3433
Mekong River Delta	1-Small	375	10	20	13	106	519	212	1273	2528
	2-Medium	258	8	15	14	26	47	148	346	862
	3-Large	87	6	31	11	8	9	88	72	312
	TOTAL	720	24	66	38	140	575	448	1691	3702
Red River Delta	1-Small	350	85	124	98	541	1411	1616	12436	16661
	2-Medium	170	79	91	111	270	172	1051	3590	5534
	3-Large	65	40	116	90	70	25	531	954	1891
	TOTAL	585	204	331	299	881	1608	3198	16980	24086
South Central Coast	1-Small	120	13	19	13	122	974	231	1828	3320
	2-Medium	49	13	18	19	41	59	135	648	982
	3-Large	44	9	19	10	10	2	80	213	387
	TOTAL	213	35	56	42	173	1035	446	2689	4689
South East	1-Small	545	224	440	197	675	3265	2636	13363	21345
	2-Medium	262	193	342	231	426	312	2101	3488	7355
	3-Large	200	182	509	133	192	62	1503	973	3754
	TOTAL	1007	599	1291	561	1293	3639	6240	17824	32454
Grand Total		2619	873	1769	1145	2549	7197	10622	41590	68364

ELIGIBLE UNIVERSE (Panel)										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	3	0	1	3	5	0	6	0	18
	2-Medium	6	1	2	13	3	0	17	0	42
	3-Large	9	1	2	15	1	0	13	0	41
	Region Total	18	2	5	31	9	0	36	0	101
Mekong River Delta	1-Small	3	0	0	0	1	0	0	0	4
	2-Medium	19	1	1	0	8	0	15	0	44
	3-Large	13	2	11	3	1	0	11	0	41
	Region Total	35	3	12	3	10	0	26	0	89
Red River Delta	1-Small	7	1	2	0	5	0	23	0	38
	2-Medium	8	4	4	5	6	0	60	0	87
	3-Large	16	4	24	26	7	0	82	0	159
	Region Total	31	9	30	31	18	0	165	0	284
South Central Coast	1-Small	8	2	2	11	6	0	24	0	53
	2-Medium	16	7	4	14	16	0	76	0	133
	3-Large	25	16	32	20	9	0	91	0	193
	Region Total	49	25	38	45	31	0	191	0	379
South East	1-Small	1	0	0	1	5	0	5	0	12
	2-Medium	9	0	1	0	4	0	16	0	30
	3-Large	11	2	6	2	2	0	18	0	41
	Region Total	21	2	7	3	11	0	39	0	83
Grand Total		154	41	92	113	79	0	457	0	936

9. The enumerated establishments were then used as the frame for the selection of a sample with the aim of obtaining interviews at 1050 establishments with five or more employees

10. The quality of the frame was assessed at the onset of the project through calls to a random subset of firms and local contractor knowledge. The sample frame was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc. Due to response rate and ineligibility issues, additional sample had to be extracted by the World Bank in order to obtain enough eligible contacts and meet the sample targets.

11. 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 23% (734 out of 3131 establishments).<sup>4</sup> Breaking down by industry, the following numbers of establishments were surveyed:

15 (Food)	127
17 (Textiles)	120
18 (Garments)	120
26 (Non-metallic mineral products)	123

<sup>4</sup> Appendix B shows the tabulations for the Philippines sample of registered firms of response codes that are classified as eligible and non-eligible.

28(Metal & Fabrication)	122
Other manufacturing	196
Retail & IT	128
Other services	117

**Local Agency team involved in the study:**

Local Agency	Name: Mekong Economics Ltd. Location: Hanoi, Vietnam Membership of international organization: N/A Activities since: 2000
Name of Project Manager	Mr. Adam McCarty
Name and position of other key persons of the project: Local Survey Implementation Team and corresponding supervisor and enumerator codes:	Ms. Tran Ngoc Diep Vietnam Country Project Coordinator Ms. Vu Thi Linh Vietnam Fieldwork Manager Ms. Doan Hai Chau Lead Supervisor Ms. Nguyen Thi Minh Hieu Data entry Specialist
Other staff involved:	59 Enumerators and 11 screeners

**Sample Frame:**

Characteristics of sample frame used	Variables: Name of establishment, address, activity, telephone number, number of employees
Source:	Vietnam General Statistics Office
Year:	
Comments on the quality of sample frame:	Good quality and coverage, and contained all the variables needed to construct a sample frame for the enterprise survey. Some addresses, telephone number outdated. There are some firms missing contact information.
Year and organism who conducted the last economic census	General Statistics Office, Ministry of Planning and Investment, Vietnam
Other sources for companies statistics	Oral/face-to-face interview

**Sectors included in the sample:**

Original Sectors	Manufacturing: 15 - 37 Retail: 52 Services:
------------------	--

	45,50,51,55,60,61,62,63,64,72
--	-------------------------------

### Sample:

Comments/ problems on sectors and regions selected in the sample	On sectors: Second preference/replacements for large establishments were sometimes not available in the sample.  Region codes: A = Red River Delta B = North Central Coast C = Mekong River Delta D = South Central Coast E = South East
Comments on the response rate	The overall response rate of establishments was medium.
Comments on the sample design:	
Other comments:	None

### Fieldwork:

Date of Fieldwork	June 2009- January 2010
Locations	Hanoi, Ha Tay, Hai Duong, Hai Phong, Thanh Hoa, Nghe An, Khanh Hoa, Da Nang, Binh Duong, Dong Nai, Ho Chi Minh, Long An, Tien Giang, Can Tho
Interview number	Manufacturing: 808 Services: 128 Other Services: 117
Problems found during fieldwork	Long distance to enterprises which are located outside the main urban areas, and difficulty to find the exact location or contact details of some establishments since these establishments either have moved their offices or contact details provided were not correct. Some individuals were unwilling to provide income, sales, and expenditure related information.
Other observations:	None

### 4. Data Base Structure:

12. The structure of the data base reflects the fact that 3 different versions of the questionnaire were used. The basic questionnaire, the Core Module, includes all common questions asked to all establishments from all sectors (manufacturing, services and IT). The second expanded variation, the Manufacturing Questionnaire, is built upon the Core Module and adds some specific questions relevant to the sector. The third expanded variation, the Services Questionnaire, is also built upon the Core Module and adds to the

core specific questions relevant to either retail or IT. Each variation of the questionnaire is identified by the index variable, *a0*.

13. 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. Variable names preceded by a prefix “EA” indicate questions specific to East Asia and Pacific and, 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.

14. 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.

In addition, establishments that were identified as having been interviewed in 2005 can be matched to the 2005 dataset using the identifier *idstd2005*. A variable *panel* describes which establishments have been interviewed in what years.

15. 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.

16. All of the following variables contain information from the sampling frame and were defined with the sampling design. They may not coincide with the reality of individual establishments as sample frames may contain inaccurate 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 (oblasts)

- a6a*: coded using the same standard for small, medium, and large establishments as defined above. The code -9 was used to indicate units for which size was undetermined in the sample frame.

- a4a*: coded using ISIC codes for the chosen industries for stratification. These codes include most manufacturing industries (15 to 37), retail (52), and (45, 50, 51, 55, 60, 63, 72) for Other Services.

17. The surveys were implemented following a 2 stage procedure. In the first stage a screener questionnaire was applied over the phone to determine eligibility and to make appointments; in the second stage, a face-to-face interview took place with the Manager/Owner/Director of each establishment. 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.

18. Note that there are additional variables for location (*a3x*) and 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.

19. Variable *a3x* indicates the actual location of the establishment. There may be divergences between the location in the sampling frame and the actual location, as establishments may be listed in one place but the actual physical location is in another place.

20. Variables *l1*, *l6* and *l8* 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.

## **5. Universe Estimates**

21. The enumerated totals were adjusted to take account of the establishments found to be ineligible when interviews were attempted. Then ratios of the total numbers of blocks of each type to the totals enumerated were formed. Those ratios were then applied to the eligible establishments enumerated to provide universe estimates.

22. Appendix C shows the overall estimates of the numbers of establishments in Vietnam based on the sample frame.

23. 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.

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

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

*w\_strict*.

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

26. 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 *w\_median*.

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

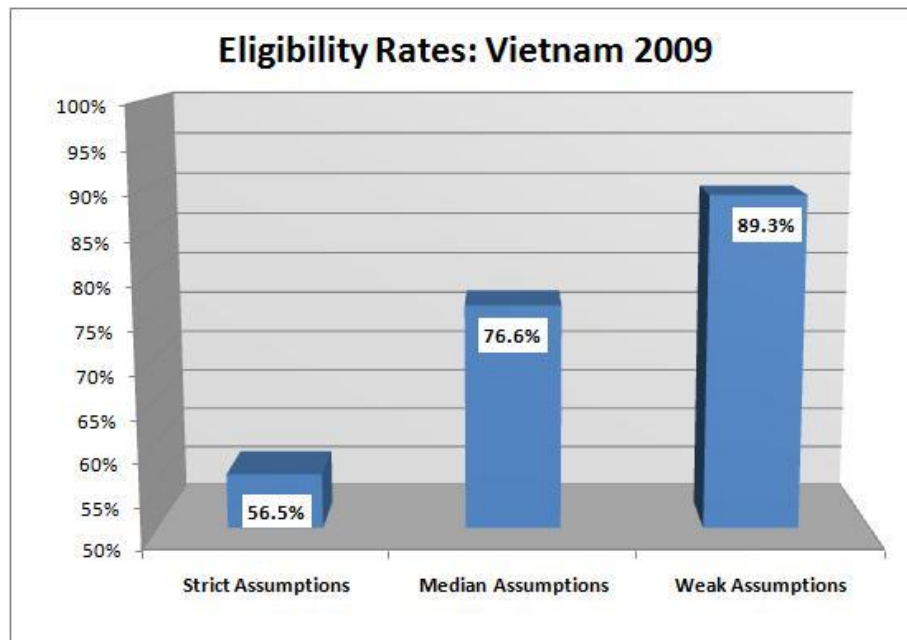
27. Weak assumption: in addition to the establishments included in points a and b, all establishments for which it was not possible to contact or that refused the screening questionnaire are assumed eligible. This definition includes as eligible establishments with dead or out of service phone lines, 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 *w\_weak*.

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

28. The indicators computed for the Enterprise Survey 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.



29. 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 C shows the universe estimates of the numbers of registered establishments.

30. 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.

## 6. Weights

31. 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.)<sup>5</sup>

<sup>5</sup> This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

32. 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, establishments with less than 5 employees, no reply after having called in different days of the week and in different business hours, out of order, no tone in the phone line, answering machine, fax line, 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.

33. Appendix D shows the cell weights for registered establishments in Vietnam.

## **7. Appropriate use of the weights**

34. 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.

35. 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 has the advantage of providing an estimate that is independent of the sample design. This latter point may be quite relevant for the Enterprise Surveys 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.)<sup>6</sup>

36. 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.<sup>7</sup> 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.

---

<sup>6</sup> 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.

<sup>7</sup> 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.

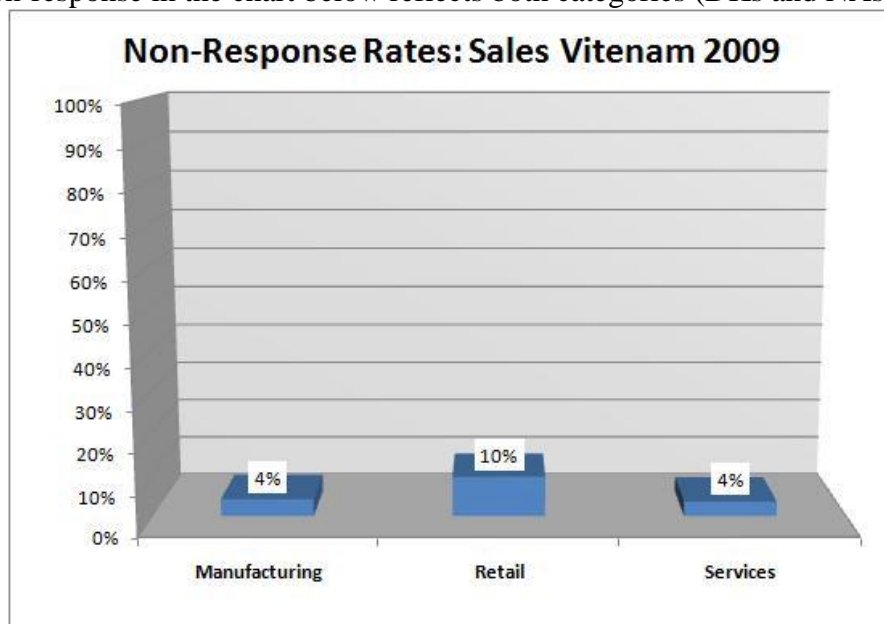
## 8. Non-response

37. 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.

38. Item non-response was addressed by two strategies:

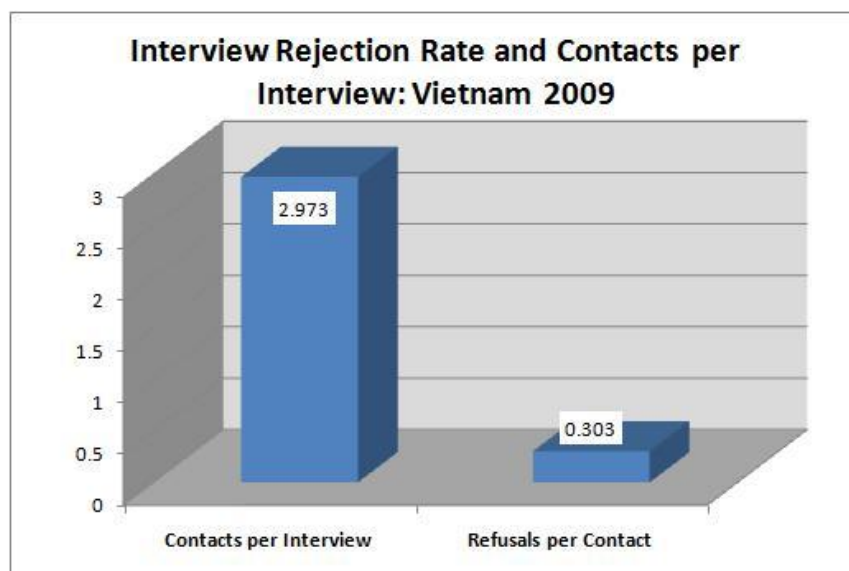
a- For sensitive questions that may generate negative reactions from the respondent, such as corruption or tax evasion, enumerators were instructed to collect the refusal to respond as a different option from don't know (-7).

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 the coding utilized in this dataset does not allow us to differentiate between "Don't know" and "refuse to answer", thus the non-response in the chart below reflects both categories (DKs and NAs).



39. 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. Further research is needed on survey non-response in the Enterprise Surveys regarding potential introduction of bias.

40. As the following graph shows, the number of contacted establishments per realized interview was 2.97. 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 number of rejections per contact was 0.303.



41. 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, 1977.

Deaton, Angus, The Analysis of Household Surveys, 1998.

Levy, Paul S. and Stanley Lemeshow, Sampling of Populations: Methods and Applications, 1999.

Lohr, Sharon L. Sampling: Design and Techniques, 1999.

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

## Appendix A

### Questionnaires:

Problems for the understanding of questions (write question number)	Missing of VN_G2, VN_G3, VNG_4. Then the information for those questions was updated during the quality control process (call-back). The survey manager and supervisors then allocated additional time for training to resolve any remaining understanding issues.
Problems found in the navigability of questionnaires (for example skip patterns)	No problems
Comments on questionnaire length:	A number of respondents commented that the questionnaire was too long, although in most cases the questionnaire was completed taking under one hour.
Suggestions or other comments on the questionnaire:	Make question wording shorter and more concise

### Database:

Comments on the data entry program	Data entry program chosen: CSPro The software program was in general easy to use
Comments on the data cleaning	Checking for data errors and inconsistencies was conducted by MKE and a quality control report and list of corrections was provided to the data entry supervisor.

### Country Situation:

General aspects of the economic, political or Social situation in your territory that could affect the results of the survey:	Vietnam is a one party communist state and the political and social situation in the country is currently stable. Vietnam is classified as a least Development Country (LDC), with an agricultural based economy and over 80% of the population residing in rural areas. Non primary industries are at a relatively early stage of development. As a result there were few large scale establishments in the manufacturing and service sector under scope of the survey.
---	--

Relevant local events occurred during fieldwork:	Vietnam Independence Day on 2 <sup>nd</sup> September 2009 Vietnam New Year Festival on 1 January 2010
Other aspects:	

## Appendix B

### Status Codes Panel:

	Panel Establishments	
Eligibles	1. Eligible establishment (Correct name and address)	386
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	5
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	7
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	19
Ineligibles	5. The establishment has less than 5 permanent full time employees	2
	6. The firm discontinued businesses	48
	7. Not a business: private household	9
	8. Ineligible activity: education, agriculture, finances, governments...	25
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	36
	92. Line out of order	44
	93. No tone	31
	10. Answering machine	1
	11. Fax line - data line	2
	12. Wrong address/ moved away and could not get the new references	24
	13. Refuses to answer the screener	210
	<b>14. In process</b> (the establishment is being called/ is being contacted - previous to ask the screener)	<b>84</b>
	151. Out of target - outside the covered regions, firm moved abroad	3
	152. Out of target - firm moved abroad	0
	<b>Total Contacts used</b>	<b>936</b>

### Response Outcomes Panel:

Panel Establishments	
Complete interviews (Total)	357
Incomplete interviews	0
Elegible in process	14
Refusals	46
Out of target	84
Impossible to contact	138
Ineligible - coop.	87
Refusal to the Screener	210
	936



### Status Codes Fresh:

	Fresh Establishments	
Eligibles	1. Eligible establishment (Correct name and address)	828
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	2
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	0
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	3
Ineligibles	5. The establishment has less than 5 permanent full time employees	8
	6. The firm discontinued businesses	40
	7. Not a business: private household	0
	8. Ineligible activity: education, agriculture, finances, governments...	17
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	48
	92. Line out of order	76
	93. No tone	67
	10. Answering machine	2
	11. Fax line - data line	14
	12. Wrong address/ moved away and could not get the new references	208
	13. Refuses to answer the screener	612
	<b>14. In process</b> (the establishment is being called/ is being contacted - previous to ask the screener)	<b>263</b>
	151. Out of target - outside the covered regions, firm moved abroad	7
	152. Out of target - firm moved abroad	0
	<b>Total Contacts used</b>	<b>2195</b>

### Response Outcomes Fresh:

Fresh Establishments	
Complete interviews (Total)	696
Incomplete interviews	0
Elegible in process	56
Refusals	81
Out of target	65
Impossible to contact	415
Ineligible - coop.	270
Refusal to the Screener	612
	2195



## Appendix C

### Universe Estimates, Vietnam:

ELIGIBLE UNIVERSE (Fresh)										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	46	3	8	35	41	301	129	1483	2046
	2-Medium	25	5	6	116	15	35	122	768	1092
	3-Large	23	3	11	54	6	4	39	155	295
	TOTAL	94	11	25	205	62	340	290	2406	3433
Mekong River Delta	1-Small	375	10	20	13	106	519	212	1273	2528
	2-Medium	258	8	15	14	26	47	148	346	862
	3-Large	87	6	31	11	8	9	88	72	312
	TOTAL	720	24	66	38	140	575	448	1691	3702
Red River Delta	1-Small	350	85	124	98	541	1411	1616	12436	16661
	2-Medium	170	79	91	111	270	172	1051	3590	5534
	3-Large	65	40	116	90	70	25	531	954	1891
	TOTAL	585	204	331	299	881	1608	3198	16980	24086
South Central Coast	1-Small	120	13	19	13	122	974	231	1828	3320
	2-Medium	49	13	18	19	41	59	135	648	982
	3-Large	44	9	19	10	10	2	80	213	387
	TOTAL	213	35	56	42	173	1035	446	2689	4689
South East	1-Small	545	224	440	197	675	3265	2636	13363	21345
	2-Medium	262	193	342	231	426	312	2101	3488	7355
	3-Large	200	182	509	133	192	62	1503	973	3754
	TOTAL	1007	599	1291	561	1293	3639	6240	17824	32454
Grand Total		2619	873	1769	1145	2549	7197	10622	41590	68364

ELIGIBLE UNIVERSE (Panel)										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	3	0	1	3	5	0	6	0	18
	2-Medium	6	1	2	13	3	0	17	0	42
	3-Large	9	1	2	15	1	0	13	0	41
	Region Total	18	2	5	31	9	0	36	0	101
Mekong River Delta	1-Small	3	0	0	0	1	0	0	0	4
	2-Medium	19	1	1	0	8	0	15	0	44
	3-Large	13	2	11	3	1	0	11	0	41
	Region Total	35	3	12	3	10	0	26	0	89
Red River Delta	1-Small	7	1	2	0	5	0	23	0	38
	2-Medium	8	4	4	5	6	0	60	0	87
	3-Large	16	4	24	26	7	0	82	0	159
	Region Total	31	9	30	31	18	0	165	0	284
South Central Coast	1-Small	8	2	2	11	6	0	24	0	53
	2-Medium	16	7	4	14	16	0	76	0	133
	3-Large	25	16	32	20	9	0	91	0	193
	Region Total	49	25	38	45	31	0	191	0	379
South East	1-Small	1	0	0	1	5	0	5	0	12
	2-Medium	9	0	1	0	4	0	16	0	30
	3-Large	11	2	6	2	2	0	18	0	41
	Region Total	21	2	7	3	11	0	39	0	83
Grand Total		154	41	92	113	79	0	457	0	936

## Appendix D

### Strict Cell Weights Vietnam:

STRICT Weights									
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services
North Central Coast	1-Small	14.9	N/A	N/A	6.7	11.7	23.7	30.3	69.7
	2-Medium	11.3	N/A	1.9	6.6	4.0	1.6	17.7	57.5
	3-Large	5.0	2.3	1.4	6.5	N/A	N/A	5.6	13.9
Mekong River Delta	1-Small	43.2	3.2	N/A	1.3	4.3	13.2	N/A	56.9
	2-Medium	12.8	N/A	6.8	N/A	1.5	2.0	15.3	14.4
	3-Large	8.4	N/A	3.3	N/A	5.4	N/A	12.3	6.4
Red River Delta	1-Small	16.1	2.2	8.8	4.9	20.0	25.4	147.9	202.6
	2-Medium	8.5	2.8	3.7	7.7	9.8	5.2	34.6	146.7
	3-Large	5.5	4.7	3.0	3.1	3.7	4.3	19.1	77.8
South Central Coast	1-Small	11.0	1.0	N/A	6.4	5.9	29.6	92.2	146.2
	2-Medium	5.4	1.2	N/A	2.6	3.9	3.4	18.8	48.1
	3-Large	6.7	N/A	2.0	8.2	N/A	N/A	11.8	14.2
South East	1-Small	19.1	3.0	8.1	6.7	20.8	36.3	268.2	466.7
	2-Medium	14.0	3.3	8.8	6.7	14.6	10.1	77.7	132.0
	3-Large	8.8	3.2	8.9	4.6	13.2	6.6	45.0	56.7

### Median Cell Weights Vietnam:

MEDIAN Weights									
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services
North Central Coast	1-Small	15.9	N/A	N/A	7.2	14.4	31.7	27.5	120.2
	2-Medium	10.7	N/A	2.4	6.4	4.4	1.9	14.4	88.4
	3-Large	4.4	2.9	1.6	5.7	N/A	N/A	4.1	19.7
Mekong River Delta	1-Small	59.4	6.4	N/A	1.8	6.8	22.8	N/A	126.2
	2-Medium	15.6	N/A	11.2	N/A	2.1	3.1	16.0	28.4
	3-Large	9.5	N/A	5.1	N/A	7.1	N/A	11.8	11.8
Red River Delta	1-Small	20.3	3.9	14.9	6.2	29.2	40.4	159.0	413.7
	2-Medium	9.6	4.6	5.7	8.8	12.8	7.4	33.2	267.3
	3-Large	5.7	7.1	4.2	3.2	4.5	5.6	16.9	130.8
South Central Coast	1-Small	13.3	1.8	N/A	7.7	8.2	44.9	94.5	284.6
	2-Medium	5.8	1.8	N/A	2.8	4.9	4.6	17.2	83.6
	3-Large	6.7	N/A	2.7	8.2	N/A	N/A	10.0	22.8
South East	1-Small	29.0	6.6	16.6	10.2	36.6	69.3	346.7	1146.0
	2-Medium	18.9	6.5	16.1	9.1	23.0	17.2	89.6	289.3
	3-Large	11.0	5.8	15.1	5.8	19.1	10.4	47.9	114.6

## Weak Cell Weights Vietnam:

WEAK Weights									
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services
North Central Coast	1-Small	21.1	N/A	N/A	10.2	18.8	52.1	32.6	179.0
	2-Medium	11.6	N/A	2.8	7.3	4.6	2.6	13.8	107.0
	3-Large	4.4	2.9	1.8	6.2	N/A	N/A	3.7	22.4
Mekong River Delta	1-Small	83.9	8.9	N/A	2.8	9.5	39.8	N/A	199.5
	2-Medium	18.0	N/A	13.7	N/A	2.3	4.4	16.3	36.5
	3-Large	10.2	N/A	5.9	N/A	7.5	N/A	11.3	14.2
Red River Delta	1-Small	26.3	5.0	20.7	8.6	37.2	64.6	183.4	599.9
	2-Medium	10.1	4.7	6.4	9.8	13.3	9.7	31.1	315.0
	3-Large	5.6	6.9	4.4	3.4	4.3	6.8	14.8	144.5
South Central Coast	1-Small	17.5	2.3	N/A	10.8	10.6	72.9	110.9	419.5
	2-Medium	6.2	1.9	N/A	3.2	5.2	6.1	16.4	100.2
	3-Large	6.7	N/A	2.9	8.7	N/A	N/A	8.9	25.6
South East	1-Small	39.8	8.9	24.3	15.0	49.3	117.2	423.3	1758.9
	2-Medium	21.1	7.1	19.1	10.8	25.1	23.7	88.9	360.8
	3-Large	11.5	6.0	16.8	6.5	19.6	13.4	44.6	134.1

## Appendix E

### Original Sample Design, Vietnam:

*The original aim was to obtain 1050 interviews consisting of; 120 to 145 interviews within each of the 6 manufacturing sectors (15-Food, 17-Textiles, 18-Garments, 26-Non-Metalic, 28 Metal & Fabrication, and Other Manufacturing; along with 120 interviews with retail establishments; and 120 interviews with establishments in other services sectors.*

DESIGN										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	3	1	0	4	2	5	1	4	21
	2-Medium	1	0	1	14	1	1	1	2	22
	3-Large	1	1	0	7	0	0	0	0	10
	TOTAL	5	2	2	25	3	6	3	7	53
Mekong River Delta	1-Small	21	2	1	2	6	9	2	4	46
	2-Medium	14	1	1	2	1	1	2	1	23
	3-Large	5	3	1	1	0	0	1	0	11
	TOTAL	40	5	3	5	8	10	5	5	80
Red River Delta	1-Small	19	10	12	12	30	24	18	36	161
	2-Medium	9	7	11	14	15	3	12	10	81
	3-Large	4	10	5	11	4	0	6	3	43
	TOTAL	32	27	28	37	48	27	36	49	284
South Central Coast	1-Small	7	2	2	2	7	16	3	5	42
	2-Medium	3	1	2	2	2	1	2	2	15
	3-Large	2	2	1	1	1	0	1	1	9
	TOTAL	12	5	5	5	10	17	5	8	66
South East	1-Small	30	36	31	24	37	54	30	39	281
	2-Medium	15	28	27	28	23	5	24	10	160
	3-Large	11	42	25	16	11	1	17	3	125
	TOTAL	56	106	82	69	71	61	70	51	566
Grand Total		145	120	145	140	140	120	120	120	1050

*A total of 3131 eligible establishments were enumerated. The World Bank drew sample and issued replacements. The implementing agency began fieldwork by first contacting and interviewing as many of the 936 establishments with contacts in the Panel frame as possible. This resulted in a total of 358 Panel interviews with many in the "Other Manufacturing" sector. The sample design was adjusted to account for the panel interviews. The fresh contacts were then issued and given preferences from 1 to 8. The implementing agency was instructed to exhaust contacts with preference-1 establishments before moving on to subsequent preferences. For some cells, the contacts were exhausted before completing the sample. In most cases the team did not have to go beyond preference 4 to complete the sample.*



## Completed Interviews, Vietnam:

Realized Complete Interviews										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	2			3	2	5	4	8	24
	2-Medium	2		2	14	3	12	9	7	49
	3-Large	5	1	6	8			11	7	38
	Region Total	9	1	8	25	5	17	24	22	111
Mekong River Delta	1-Small	4	1		4	10	11		6	36
	2-Medium	13		1		10	9	9	9	51
	3-Large	8		5		1		8	5	27
	Region Total	25	1	6	4	21	20	17	20	114
Red River Delta	1-Small	11	14	5	9	12	17	8	18	94
	2-Medium	14	14	12	9	17	14	31	10	121
	3-Large	10	5	23	22	14	3	34	6	117
	Region Total	35	33	40	40	43	34	73	34	332
South Central Coast	1-Small	6	5		1	10	11	2	4	39
	2-Medium	7	6		5	7	8	8	6	47
	3-Large	6		6	1			9	8	30
	Region Total	19	11	6	7	17	19	19	18	116
South East	1-Small	12	22	16	11	12	23	6	7	109
	2-Medium	11	24	16	18	15	11	23	9	127
	3-Large	16	28	28	18	9	4	34	7	144
	Region Total	39	74	60	47	36	38	63	23	380
Grand Total		127	120	120	123	122	128	196	117	1053

Realized Complete Interviews - Fresh										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small				2		5		8	15
	2-Medium				7		12	4	7	30
	3-Large			5	4			2	7	18
	Region Total			5	13		17	6	22	63
Mekong River Delta	1-Small	4	1		4	10	11		6	36
	2-Medium	7				6	9	2	9	33
	3-Large	4						2	5	11
	Region Total	15	1		4	16	20	4	20	80
Red River Delta	1-Small	10	13	4	9	10	17	2	18	83
	2-Medium	10	12	10	8	14	14	5	10	83
	3-Large	3	4	7	10	7	3	5	6	45
	Region Total	23	29	21	27	31	34	12	34	211
South Central Coast	1-Small	5	5			9	11		4	34
	2-Medium	5	6		5	5	8	5	6	40
	3-Large	3							8	11
	Region Total	13	11		5	14	19	5	18	85
South East	1-Small	11	20	16	9	8	23		7	94
	2-Medium	7	23	14	12	9	11	2	9	87
	3-Large	7	21	17	9	6	4	4	7	75
	Region Total	25	64	47	30	23	38	6	23	256
Grand Total		76	105	73	79	84	128	33	117	695

Realized Complete Interviews - Panel										
Region	Size	15 Food	17 Textiles	18 Garments	26 Non-metallic	28 Metal & fabrication	52 Retail	Other Manufacturing	Other Services	TOTAL
North Central Coast	1-Small	2			1	2		4		9
	2-Medium	2		2	7	3		5		19
	3-Large	5	1	1	4			9		20
	Region Total	9	1	3	12	5		18		48
Mekong River Delta	1-Small									
	2-Medium	6		1		4		7		18
	3-Large	4		5		1		6		16
	Region Total	10		6		5		13		34
Red River Delta	1-Small	1	1	1		2		6		11
	2-Medium	4	2	2	1	3		26		38
	3-Large	7	1	16	12	7		29		72
	Region Total	12	4	19	13	12		61		121
South Central Coast	1-Small	1			1	1		2		5
	2-Medium	2				2		3		7
	3-Large	3		6	1			9		19
	Region Total	6		6	2	3		14		31
South East	1-Small	1	2		2	4		6		15
	2-Medium	4	1	2	6	6		21		40
	3-Large	9	7	11	9	3		30		69
	Region Total	14	10	13	17	13		57		124
Grand Total		51	15	47	44	38		163		358