

## **The Yemen 2013 Enterprise Surveys Data Set**

### **I. Introduction**

1. This document provides additional information on the data collected in Yemen between March 2013 and July 2014. The survey was part of the Joint World Bank/European Bank for Reconstruction and Development (EBRD)/European Investment Bank (EIB) Enterprise Survey, which is an enterprise survey whose objective is to gain an understanding of firms' perception of the environment in which they operate. This has added an important element of dynamics in the study of business environment in transition countries.

The Enterprise Surveys, through interviews with firms in the manufacturing and services sectors, capture business perceptions on the biggest obstacles to enterprise growth, the relative importance of various constraints to increasing employment and productivity, and the effects of a country's business environment on its international competitiveness. They are used to create statistically significant business environment indicators that are comparable across countries. The Enterprise Surveys 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.

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.

### **II. Sampling Structure**

2. The sample for Yemen 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.

---

<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

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.

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 manufacturing and two service industries (retail and other services).

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 6 regions (due to classifications in the sample frame, regions were defined at the governorate level) throughout Yemen. The six regional strata included were: Amanat Al-Asemah (Sana'a), Aden, Hudaydah, Hadramaut, Ibb, and Taizz.

### **III. 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. Great efforts were made to obtain the best source for these listings. However, the quality of the sample frames was not optimal and, therefore, some adjustments were needed to correct for the presence of ineligible units. These adjustments are reflected in the weights computation (*see below*).

8. The Gallup Organization and the Yemen Polling Center (YPC) were hired to implement the Yemen 2013 enterprise survey.

9. For Yemen, two sample frames were used. The first was supplied by the World Bank and consists of enterprises interviewed in Yemen in 2010. The World Bank required that attempts should be made to re-interview establishments responding to the Yemen 2010 Enterprise Survey where they met eligibility criteria. This sample is referred to as the Panel. The second sample frame, here referred to as the Fresh sample, was obtained from the Central Statistics Office, 2010 Establishment Census, with updates and validation provided by YPC.

The database contained the following information

- Coverage;
- Up to datedness;- Availability of detailed stratification variables;
- Contact name(s).

Counts from the sample frame are shown below.

## Sample Frames

Panel (Source: World Bank Enterprise Survey, 2010)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	31	6	24	61
	Medium	20	13	14	47
	Large	10	1	12	23
REGIONAL TOTAL		61	20	50	131
<b>Aden</b>	Small	23	17	10	50
	Medium	12	3	14	29
	Large	5	1	3	9
REGIONAL TOTAL		40	21	27	88
<b>Hudaydah</b>	Small	39	16	12	67
	Medium	10	1	8	19
	Large	5	0	0	5
REGIONAL TOTAL		54	17	20	91
<b>Hadramaut</b>	Small	13	7	6	26
	Medium	9	0	3	12
	Large	6	1	0	7
REGIONAL TOTAL		28	8	9	45
<b>Ibb</b>	Small	19	6	9	34
	Medium	3	0	3	6
	Large	0	0	0	0
REGIONAL TOTAL		22	6	12	40
<b>Taizz</b>	Small	33	14	11	58
	Medium	7	0	9	16
	Large	6	0	2	8
REGIONAL TOTAL		46	14	22	82
<b>GRAND TOTAL</b>		<b>251</b>	<b>86</b>	<b>140</b>	<b>477</b>

Fresh (Source: Central Statistics Office, 2010 Establishment Census, with updates provided by contractor)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1,510	791	2,296	4,597
	Medium	73	44	180	297
	Large	11	5	13	29
REGIONAL TOTAL		1,594	840	2,489	4,923
<b>Aden</b>	Small	346	186	659	1,191
	Medium	19	13	50	82
	Large	9	0	11	20
REGIONAL TOTAL		374	199	720	1,293
<b>Hudaydah</b>	Small	611	228	841	1,680
	Medium	26	13	44	83
	Large	12	1	3	16
REGIONAL TOTAL		649	242	888	1,779
<b>Hadramaut</b>	Small	366	104	530	1,000
	Medium	24	13	21	58
	Large	7	0	3	10
REGIONAL TOTAL		397	117	554	1,068
<b>Ibb</b>	Small	211	170	356	737
	Medium	4	12	11	27
	Large	0	0	0	0
REGIONAL TOTAL		215	182	367	764
<b>Taizz</b>	Small	519	239	776	1,534
	Medium	18	12	25	55
	Large	17	0	3	20
REGIONAL TOTAL		554	251	804	1,609
<b>GRAND TOTAL</b>		<b>3,783</b>	<b>1,831</b>	<b>5,822</b>	<b>11,436</b>

10. The enumerated establishments were then used as the frame for the selection of a sample with the aim of obtaining interviews at 360 establishments with five or more employees.

11. The quality of the frame was assessed at the onset of the project through visits 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.

12. 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 14.5% (165 out of 1,141 establishments).<sup>4</sup> Breaking down by stratified industries, the following sample targets were achieved (using a4a and a6a):

Achieved Sample:

#### COMPLETED PANEL

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	10	1	8	19
	Medium	6	3	5	14
	Large	4	0	7	11
REGIONAL TOTAL		20	4	20	44
<b>Aden</b>	Small	4	5	2	11
	Medium	3	1	5	9
	Large	0	0	1	1
REGIONAL TOTAL		7	6	8	21
<b>Hudaydah</b>	Small	4	5	5	14
	Medium	2	0	2	4
	Large	0	0	0	0
REGIONAL TOTAL		6	5	7	18
<b>Hadramaut</b>	Small	6	2	1	9
	Medium	5	0	1	6
	Large	4	0	0	4
REGIONAL TOTAL		15	2	2	19
<b>Ibb</b>	Small	6	1	2	9
	Medium	2	0	1	3
	Large	0	0	0	0
REGIONAL TOTAL		8	1	3	12
<b>Taizz</b>	Small	3	4	5	12
	Medium	5	0	3	8
	Large	3	0	0	3
REGIONAL TOTAL		11	4	8	23
<b>GRAND TOTAL</b>		<b>67</b>	<b>22</b>	<b>48</b>	<b>137</b>

<sup>4</sup> Based on out of target contacts and impossible to contact establishments

### COMPLETED FRESH

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	9	41	21	71
	Medium	9	9	11	29
	Large	3	0	2	5
REGIONAL TOTAL		21	50	34	105
<b>Aden</b>	Small	3	4	0	7
	Medium	6	2	6	14
	Large	1	0	1	2
REGIONAL TOTAL		10	6	7	23
<b>Hudaydah</b>	Small	1	5	1	7
	Medium	6	2	6	14
	Large	2	0	0	2
REGIONAL TOTAL		9	7	7	23
<b>Hadramaut</b>	Small	4	7	4	15
	Medium	7	3	8	18
	Large	1	0	0	1
REGIONAL TOTAL		12	10	12	34
<b>Ibb</b>	Small	2	4	0	6
	Medium	1	0	1	2
	Large	0	0	0	0
REGIONAL TOTAL		3	4	1	8
<b>Taizz</b>	Small	2	1	3	6
	Medium	5	1	9	15
	Large	1	0	1	2
REGIONAL TOTAL		8	2	13	23
<b>GRAND TOTAL</b>		<b>63</b>	<b>79</b>	<b>74</b>	<b>216</b>

### IV. Data Base Structure:

13. 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. The second expanded variation, the Manufacturing Questionnaire, is built upon the Core Module and adds some specific questions relevant to manufacturing sectors. The third expanded variation, the Retail Questionnaire, is also built upon the Core Module and adds to the core specific questions relevant to retail firms. Each variation of the questionnaire is identified by the index variable, *a0*.

14. 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 “MNA” indicate questions specific to the Middle East and North Africa region, 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.

15. There are 3 establishment identifiers, *idstd*, *phoneid* and *id*. The first is a global unique identifier. The second two are country unique identifiers. 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.

16. 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, whereas the latter gives the actual establishment’s industry classification (four digit code) in the sample frame.

17. 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 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. The code -9 was used to indicate units for which size was undetermined in the sample frame.

- a4a*: coded using ISIC Rev 3.1 codes for the chosen industries for stratification. These codes include most manufacturing industries (15 to 37), retail (52), and (45, 50, 51, 55, 60-64, 72) for other services.

18. 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. 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.

19. 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.

20. 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.

21. 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.

22. Variables *a17x* gives interviewer comments, including problems that occurred during an interview and extraordinary circumstances which could influence results. Please note that sometimes this variable is removed due to privacy issues.

## **V. Universe Estimates**

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

24. Appendix B shows the overall estimates of the numbers of establishments in Yemen based on the sample frame.

25. 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.

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

27. 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***

28. 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*.

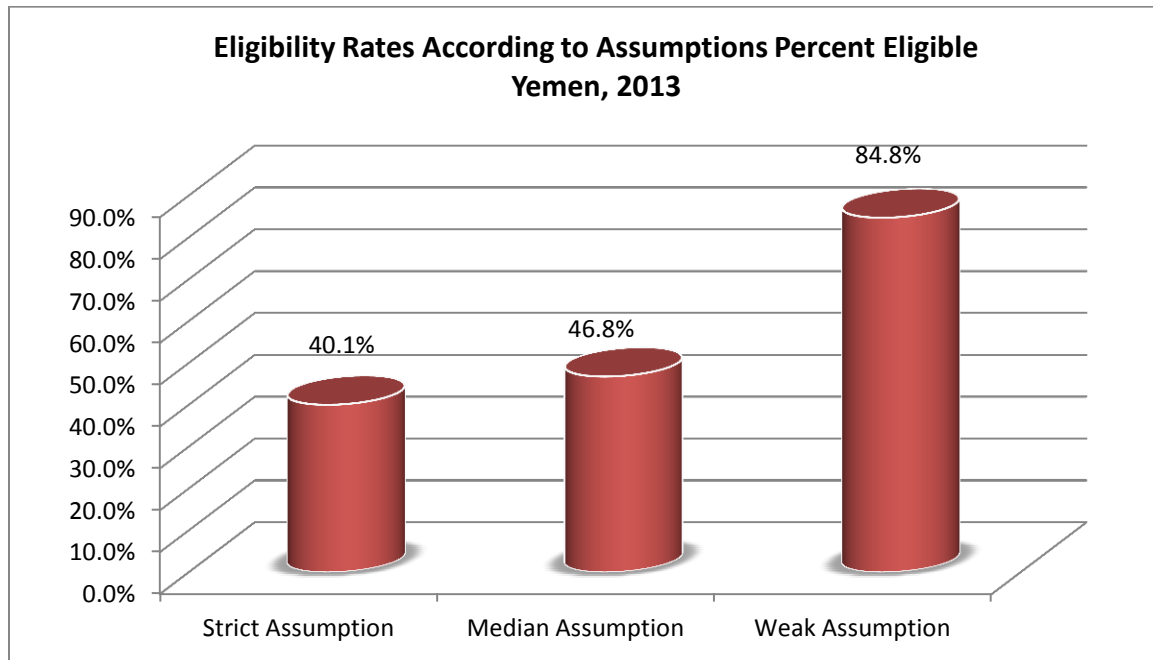
***Median eligibility = (Sum of the firms with codes 1,2,3,4,16,10,11, & 13) / Total***

29. 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 1,2,3,4,16,91,92,93,10,11,12,&13) / Total***

30. 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.



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

32. 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**

33. 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>

34. 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, no tone on the phone line, answering machine, or fax line<sup>6</sup>, 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.

35. Appendix C shows the cell weights for registered establishments in Yemen.

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

36. 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.

37. 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 a 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>7</sup>

38. 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

---

<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.

<sup>6</sup> For the surveys that implemented a screener over the phone.

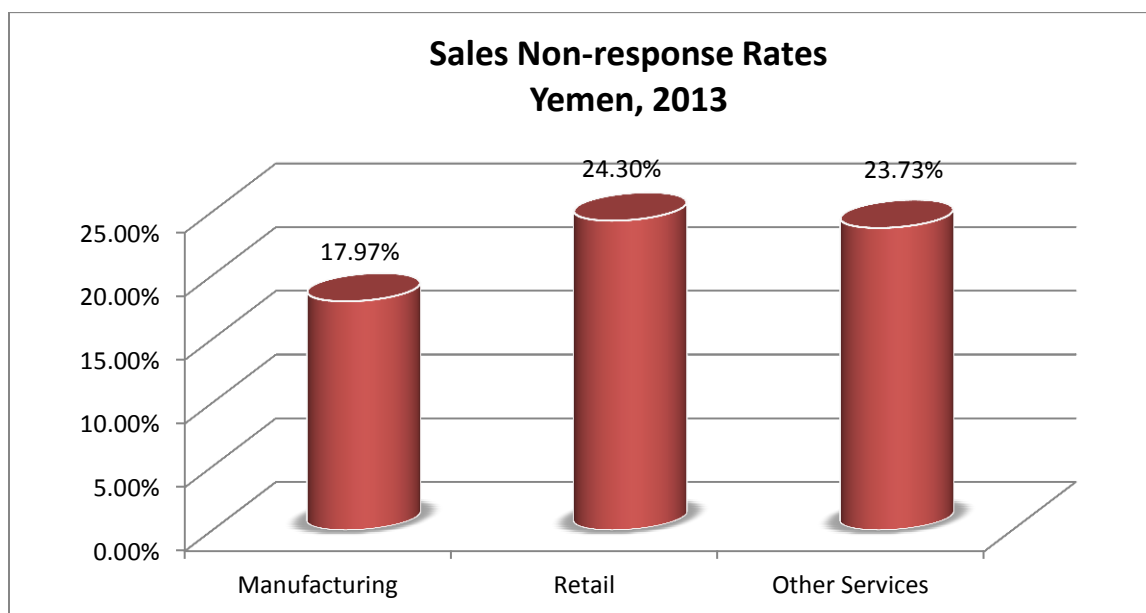
<sup>7</sup> Note that weighted OLS in Stata using the command `regress` with the option of weights will estimate wrong standard errors. Using the Stata survey specific commands `svy` will provide appropriate standard errors.

that would be expected if the whole population were observed.<sup>8</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.

## VIII. Non-response

39. 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.

40. 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 (-8).
  - 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).

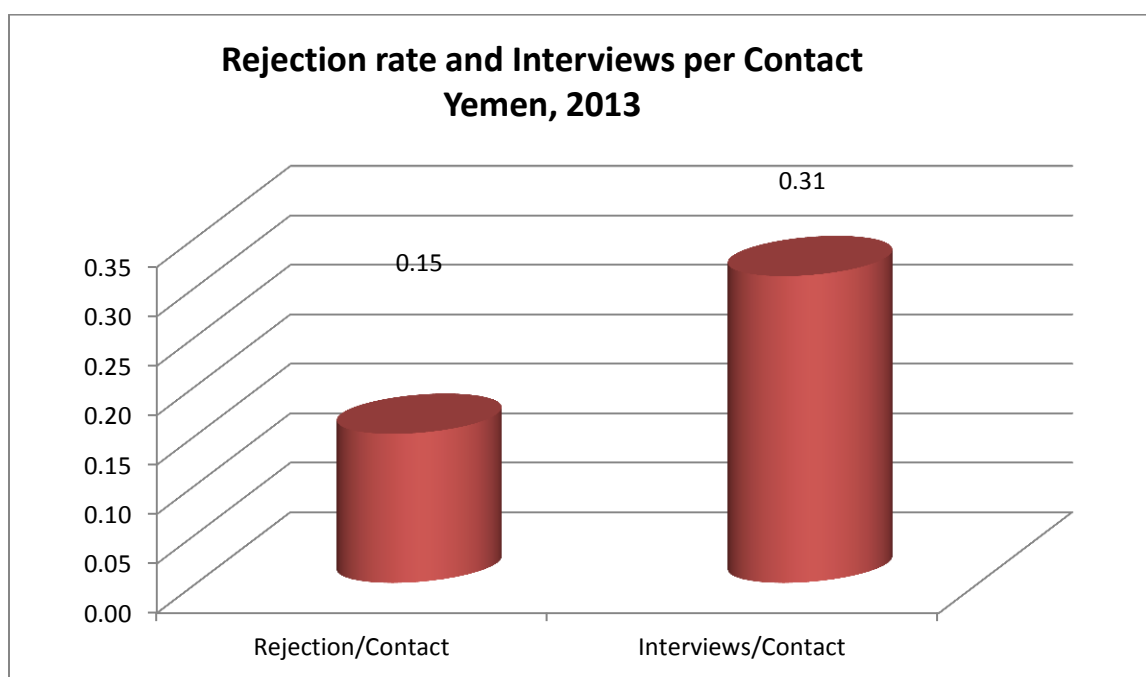


41. Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) was suggested for interview. Survey non-response did occur but substitutions were made in order to potentially achieve strata-

<sup>8</sup> The use of 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.

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

42. As the following graph shows, the number of realized interviews per contacted establishment was 0.31.<sup>9</sup> 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 0.15.



43. Details on the rejection rate, eligibility rate, and item non-response are available at the strata level. 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 Yemen. All Enterprise Surveys suffer from these shortcomings, but in very few cases they have been made explicit.

---

<sup>9</sup> The estimate is based on the total number of firms contacted including ineligible establishments.

**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

### Status Codes:

Eligibles	1. Eligible establishment (Correct name and address)	127
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	1
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	46
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	256
	16. Panel firm - now less than five employees	27
Ineligible	5. The establishment has less than 5 permanent full time employees	53
	616 The firm discontinued businesses - (Establishment went bankrupt)	19
	618 The firm discontinued businesses - (Original establishment disappeared and is now a different firm)	2
	619 The firm discontinued businesses - (Establishment was bought out by another firm)	2
	620 The firm discontinued businesses - (It was impossible to determine for what reason)	1
	621 The firm discontinued businesses - (Other: SPECIFY in COMMENTS)	5
	7. Not a business: private household	63
	8. Ineligible activity: education, agriculture, finances, governments...	21
Unobtainable	91. No reply ( <i>after having called in different days of the week and in different business hours</i> )	73
	92. Line out of order	117
	93. No tone	39
	94. Phone number does not exist	4
	10. Answering machine	0
	11. Fax line - data line	1
	12. Wrong address/ moved away and could not get the new references	19
	121. Wrong address/wrong name moved away and could not get the new references	181
	13. Refuses to answer the screener	76
	14. In process ( <i>the establishment is being called/ is being contacted - previous to ask the screener</i> )	0
	141. In process ( <i>the establishment is being called/ is being contacted - previous to ask the screener</i> ) - other preferences being contacted (PANEL ONLY)	0
Out of target	151. Out of target - outside the covered regions, firm moved abroad	0
	152. Out of target - firm moved abroad	3
	153. Out of target - Not registered with SAT	5
Total		1141

## Response Outcomes Total:

	Yemen		COMPLETE
<b>Sample Target</b>	<b>360</b>		353
1. Complete interviews <b>(Total)</b>	138	ELIGIBLES (Status Codes)	
6. Completed, eligible but refused to answer innovation	16		
2. Incomplete interviews	0		
4. Eligible in process	8		
3. Refusals	96		
5. Complete interviews with innovation (Total)	200		
Ineligible	165		
Unobtainable	434		
Out of Target	8		
<b>(Screener) In Process</b>	<b>0</b>		
Refusal to the Screener	76		
<b>Total</b>	<b>1141</b>		

<b>Response rate</b>	<b>66.3%</b>
<b>Ineligible</b>	<b>14.5%</b>
<b>Impossible to contact</b>	<b>38.0%</b>
<b>Out of target</b>	<b>0.7%</b>

## Appendix B

### Universe of Inference Estimate, Yemen

(Source: Central Statistics Office, Establishment Census):

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1,526	802	2,467	4,795
	Medium	110	58	212	380
	Large	16	5	32	53
REGIONAL TOTAL		1,652	865	2,711	5,228
<b>Aden</b>	Small	340	204	739	1,283
	Medium	28	16	83	127
	Large	15	0	20	35
REGIONAL TOTAL		383	220	842	1,445
<b>Hudaydah</b>	Small	642	243	937	1,822
	Medium	23	17	51	91
	Large	15	2	8	25
REGIONAL TOTAL		680	262	996	1,938
<b>Hadramaut</b>	Small	367	114	558	1,039
	Medium	37	15	40	92
	Large	10	0	6	16
REGIONAL TOTAL		414	129	604	1,147
<b>Ibb</b>	Small	239	177	380	796
	Medium	6	12	19	37
	Large	0	0	1	1
REGIONAL TOTAL		245	189	400	834
<b>Taizz</b>	Small	546	253	538	1,337
	Medium	32	13	11	56
	Large	23	0	1	24
REGIONAL TOTAL		601	266	550	1,417
<b>GRAND TOTAL</b>		<b>3,975</b>	<b>1,931</b>	<b>6,103</b>	<b>12,009</b>



## Amended Universe<sup>10</sup>

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1,541	802	2,467	4,810
	Medium	110	58	212	380
	Large	21	6	32	59
REGIONAL TOTAL		1,672	866	2,711	5,249
<b>Aden</b>	Small	369	204	739	1,312
	Medium	31	16	83	130
	Large	15	1	20	36
REGIONAL TOTAL		415	221	842	1,478
<b>Hudaydah</b>	Small	650	244	937	1,831
	Medium	36	17	52	105
	Large	17	2	8	27
REGIONAL TOTAL		703	263	997	1,963
<b>Hadramaut</b>	Small	379	114	558	1,051
	Medium	37	15	40	92
	Large	13	1	6	20
REGIONAL TOTAL		429	130	604	1,163
<b>Ibb</b>	Small	239	177	380	796
	Medium	7	12	19	38
	Large	0	0	1	1
REGIONAL TOTAL		246	189	400	835
<b>Taizz</b>	Small	552	253	787	1,592
	Medium	32	13	34	79
	Large	23	0	5	28
REGIONAL TOTAL		607	266	826	1,699
<b>GRAND TOTAL</b>		<b>4,072</b>	<b>1,935</b>	<b>6,380</b>	<b>12,387</b>

<sup>10</sup> Cells adjusted upwards when sampling frame includes more contacts than indicated in Universe estimate. These figures are utilized for weighting purposes.

## Panel Universe

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	31	6	24	61
	Medium	20	13	14	47
	Large	10	1	12	23
REGIONAL TOTAL		61	20	50	131
<b>Aden</b>	Small	23	17	10	50
	Medium	12	3	14	29
	Large	5	1	3	9
REGIONAL TOTAL		40	21	27	88
<b>Hudaydah</b>	Small	39	16	12	67
	Medium	10	1	8	19
	Large	5	0	0	5
REGIONAL TOTAL		54	17	20	91
<b>Hadramaut</b>	Small	13	7	6	26
	Medium	9	0	3	12
	Large	6	1	0	7
REGIONAL TOTAL		28	8	9	45
<b>Ibb</b>	Small	19	6	9	34
	Medium	3	0	3	6
	Large	0	0	0	0
REGIONAL TOTAL		22	6	12	40
<b>Taizz</b>	Small	33	14	11	58
	Medium	7	0	9	16
	Large	6	0	2	8
REGIONAL TOTAL		46	14	22	82
<b>GRAND TOTAL</b>		<b>251</b>	<b>86</b>	<b>140</b>	<b>477</b>

## Fresh Universe

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1,510	796	2,443	4,749
	Medium	90	45	198	333
	Large	11	5	20	36
REGIONAL TOTAL		1,611	846	2,661	5,118
<b>Aden</b>	Small	346	187	729	1,262
	Medium	19	13	69	101
	Large	10	0	17	27
REGIONAL TOTAL		375	200	815	1,390
<b>Hudaydah</b>	Small	611	228	925	1,764
	Medium	26	16	44	86
	Large	12	2	8	22
REGIONAL TOTAL		649	246	977	1,872
<b>Hadramaut</b>	Small	366	107	552	1,025
	Medium	28	15	37	80
	Large	7	0	6	13
REGIONAL TOTAL		401	122	595	1,118
<b>Ibb</b>	Small	220	171	371	762
	Medium	4	12	16	32
	Large	0	0	1	1
REGIONAL TOTAL		224	183	388	795
<b>Taizz</b>	Small	519	239	776	1,534
	Medium	25	13	25	63
	Large	17	0	3	20
REGIONAL TOTAL		561	252	804	1,617
<b>GRAND TOTAL</b>		<b>3,821</b>	<b>1,849</b>	<b>6,240</b>	<b>11,910</b>

## Appendix C

Note: due to issues of high observed weights and strata-related non-response, weights were collapsed over location, pooling observations for Aden, Hudaydah, and Taizz.

Completed counts by collapsed cells are below:

### COMPLETED PANEL (COLLAPSED)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	10	1	8	19
	Medium	6	3	5	14
	Large	4	0	7	11
REGIONAL TOTAL		20	4	20	44
<b>Aden, Hudaydah, Taizz</b>	Small	11	14	12	37
	Medium	10	1	10	21
	Large	3	0	1	4
REGIONAL TOTAL		24	15	23	62
<b>Hadramaut</b>	Small	6	2	1	9
	Medium	5	0	1	6
	Large	4	0	0	4
REGIONAL TOTAL		15	2	2	19
<b>Ibb</b>	Small	6	1	2	9
	Medium	2	0	1	3
	Large	0	0	0	0
REGIONAL TOTAL		8	1	3	12
<b>GRAND TOTAL</b>		<b>67</b>	<b>22</b>	<b>48</b>	<b>137</b>

**COMPLETED FRESH (COLLAPSED)**

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	9	41	21	71
	Medium	9	9	11	29
	Large	3	0	2	5
REGIONAL TOTAL		21	50	34	105
<b>Aden, Hudaydah, Taizz</b>	Small	6	10	4	20
	Medium	17	5	21	43
	Large	4	0	2	6
REGIONAL TOTAL		27	15	27	69
<b>Hadramaut</b>	Small	4	7	4	15
	Medium	7	3	8	18
	Large	1	0	0	1
REGIONAL TOTAL		12	10	12	34
<b>Ibb</b>	Small	2	4	0	6
	Medium	1	0	1	2
	Large	0	0	0	0
REGIONAL TOTAL		3	4	1	8
<b>GRAND TOTAL</b>		<b>63</b>	<b>79</b>	<b>74</b>	<b>216</b>

## Strict Cell Weights Yemen (COLLAPSED)

Cell average, collapsed weights shown.

### STRICT WEIGHTS FRESH (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	50.71	4.04	24.69
	Medium	3.90	1.34	4.92
	Large	1.69		3.23
<b>Aden, Hudaydah, Taizz</b>	Small	153.12	28.02	265.52
	Medium	3.30	4.64	3.70
	Large	9.24		9.32
<b>Hadramaut</b>	Small	64.63	7.43	68.45
	Medium	3.64	3.13	2.96
	Large	7.53		
<b>Ibb</b>	Small	39.61	10.60	
	Medium	1.86		5.21
	Large			

### STRICT WEIGHTS PANEL (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1.50	2.47	1.36
	Medium	1.54	1.70	1.21
	Large	1.61		1.03
<b>Aden, Hudaydah, Taizz</b>	Small	4.06	1.34	1.21
	Medium	1.30	1.53	1.30
	Large	3.33		2.90
<b>Hadramaut</b>	Small	1.32	1.82	3.42
	Medium	1.05		1.63
	Large	1.22		
<b>Ibb</b>	Small	1.68	2.70	2.22
	Medium	1.00		1.41
	Large			

## Median Cell Weights Yemen (COLLAPSED)

### MEDIAN WEIGHTS FRESH (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	56.24	4.33	29.21
	Medium	4.57	1.52	6.16
	Large	2.18		4.46
<b>Aden, Hudaydah, Taizz</b>	Small	165.76	29.31	306.57
	Medium	3.78	5.13	4.52
	Large	11.67		12.55
<b>Hadramaut</b>	Small	65.67	7.30	74.18
	Medium	3.92	3.26	3.39
	Large	8.92		
<b>Ibb</b>	Small	48.03	12.41	
	Medium	2.38		7.14
	Large			

### MEDIAN WEIGHTS PANEL (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1.75	3.02	1.49
	Medium	1.76	2.04	1.31
	Large	2.26		1.37
<b>Aden, Hudaydah, Taizz</b>	Small	5.10	1.77	1.44
	Medium	1.60	1.98	1.52
	Large	5.07		4.20
<b>Hadramaut</b>	Small	1.41	2.04	3.46
	Medium	1.10		1.62
	Large	1.57		
<b>Ibb</b>	Small	1.74	2.94	2.18
	Medium	1.00		1.36
	Large			

## Weak Cell Weights Yemen (COLLAPSED)

### WEAK WEIGHTS FRESH (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	149.21	13.74	102.05
	Medium	8.83	3.51	15.68
	Large	3.19		8.57
<b>Aden, Hudaydah, Taizz</b>	Small	244.78	51.79	596.22
	Medium	4.07	6.61	6.40
	Large	9.48		13.42
<b>Hadramaut</b>	Small	82.47	10.97	122.68
	Medium	3.58	3.56	4.08
	Large	6.16		
<b>Ibb</b>	Small	103.39	31.98	
	Medium	3.73		14.73
	Large			

### WEAK WEIGHTS PANEL (COLLAPSED)

Region		Manuf.	Retail	Services
<b>Amanat Al-Asemah (Sanaa)</b>	Small	2.85	4.59	2.87
	Medium	3.01	3.25	2.62
	Large	2.60		1.85
<b>Aden, Hudaydah, Taizz</b>	Small	7.19	2.32	2.38
	Medium	2.37	2.72	2.63
	Large	5.01		4.88
<b>Hadramaut</b>	Small	1.77	2.38	5.10
	Medium	1.44		2.50
	Large	1.38		
<b>Ibb</b>	Small	2.90	4.58	4.29
	Medium	1.35		2.80
	Large			



## Appendix E

### Strict Universe Estimates Yemen

Panel and fresh cells combined

#### STRICT WEIGHTED UNIVERSE ESTIMATES (COLLAPSED)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	471	168	529	1,169
	Medium	44	17	60	122
	Large	11		14	25
REGIONAL TOTAL		527	185	603	1,316
<b>Aden, Hudaydah, Taizz</b>	Small	963	299	1,077	2,339
	Medium	69	25	91	185
	Large	47		22	68
REGIONAL TOTAL		1,080	324	1,189	2,592
<b>Hadramaut</b>	Small	266	56	277	599
	Medium	31	9	25	65
	Large	12			12
REGIONAL TOTAL		310	65	302	677
<b>Ibb</b>	Small	7	1	2	10
	Medium	2	0	1	3
	Large	0	0	0	0
REGIONAL TOTAL		9	1	3	13
<b>GRAND TOTAL</b>		<b>1,925</b>	<b>575</b>	<b>2,097</b>	<b>4,597</b>

## Median Universe Estimates Yemen

### MEDIAN WEIGHTED UNIVERSE ESTIMATES (COLLAPSED)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	523.6	180.5	625.3	1,329.4
	Medium	51.7	19.8	74.3	145.9
	Large	15.6		18.5	34.1
REGIONAL TOTAL		590.9	200.3	718.1	1,509.4
<b>Aden, Hudaydah, Taizz</b>	Small	1,050.7	317.9	1,243.5	2,612.1
	Medium	80.4	27.7	110.2	218.2
	Large	61.9		29.3	91.2
REGIONAL TOTAL		1,192.9	345.6	1,383.0	2,921.5
<b>Hadramaut</b>	Small	271.2	55.2	300.2	626.5
	Medium	32.9	9.8	28.8	71.4
	Large	15.2			15.2
REGIONAL TOTAL		319.3	64.9	328.9	713.2
<b>Ibb</b>	Small	106.5	52.6	4.4	163.4
	Medium	4.4		8.5	12.9
	Large				0.0
REGIONAL TOTAL		110.9	52.6	12.9	176.3
<b>GRAND TOTAL</b>		<b>2,214.0</b>	<b>663.4</b>	<b>2,442.9</b>	<b>5,320.4</b>

## Weak Universe Estimates Yemen

### WEAK WEIGHTED UNIVERSE ESTIMATES (COLLAPSED)

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	1,371.4	568.0	2,165.9	4,105.4
	Medium	97.5	41.4	185.6	324.5
	Large	19.9		30.1	50.0
REGIONAL TOTAL		1,488.9	609.4	2,381.6	4,479.9
<b>Aden, Hudaydah, Taizz</b>	Small	1,547.7	550.5	2,413.4	4,511.6
	Medium	92.8	35.7	160.8	289.4
	Large	52.9		31.7	84.7
REGIONAL TOTAL		1,693.5	586.2	2,605.9	4,885.6
<b>Hadramaut</b>	Small	340.5	81.5	495.8	917.8
	Medium	32.3	10.7	35.2	78.1
	Large	11.7			11.7
REGIONAL TOTAL		384.5	92.2	531.0	1,007.7
<b>Ibb</b>	Small	224.2	132.5	8.6	365.3
	Medium	6.4		17.5	24.0
	Large				0.0
REGIONAL TOTAL		230.6	132.5	26.1	389.2
<b>GRAND TOTAL</b>		<b>3,797.5</b>	<b>1,420.4</b>	<b>5,544.5</b>	<b>10,762.4</b>

## Appendix F

### Original Sample Design, Yemen:

#### FRESH DESIGN

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	8	17	20	45
	Medium	2	21	2	25
	Large	7	3	8	18
REGIONAL TOTAL		17	41	30	88
<b>Aden</b>	Small	2	2	2	6
	Medium	2	4	2	8
	Large	6	0	6	12
REGIONAL TOTAL		10	6	10	26
<b>Hudaydah</b>	Small	2	2	2	6
	Medium	2	6	2	10
	Large	7	1	2	10
REGIONAL TOTAL		11	9	6	26
<b>Hadramaut</b>	Small	2	2	2	6
	Medium	2	7	2	11
	Large	4	0	2	6
REGIONAL TOTAL		8	9	6	23
<b>Ibb</b>	Small	2	2	2	6
	Medium	2	7	2	11
	Large	0	0	0	0
REGIONAL TOTAL		4	9	4	17
<b>Taizz</b>	Small	2	2	2	6
	Medium	2	8	2	12
	Large	6	0	2	8
REGIONAL TOTAL		10	10	6	26
<b>GRAND TOTAL</b>		<b>60</b>	<b>84</b>	<b>62</b>	<b>206</b>

# **PANEL DESIGN**

Region		Manuf.	Retail	Services	Grand Total
<b>Amanat Al-Asemah (Sanaa)</b>	Small	8	6	21	35
	Medium	3	13	3	19
	Large	8	1	9	18
REGIONAL TOTAL		19	20	33	72
<b>Aden</b>	Small	2	2	2	6
	Medium	2	3	2	7
	Large	5	1	3	9
REGIONAL TOTAL		9	6	7	22
<b>Hudaydah</b>	Small	2	2	2	6
	Medium	2	1	2	5
	Large	5	0	0	5
REGIONAL TOTAL		9	3	4	16
<b>Hadramaut</b>	Small	2	2	2	6
	Medium	2	0	2	4
	Large	5	1	0	6
REGIONAL TOTAL		9	3	4	16
<b>Ibb</b>	Small	2	2	2	6
	Medium	2	0	2	4
	Large	0	0	0	0
REGIONAL TOTAL		4	2	4	10
<b>Taizz</b>	Small	2	2	2	6
	Medium	2	0	2	4
	Large	6	0	2	8
REGIONAL TOTAL		10	2	6	18
<b>GRAND TOTAL</b>		<b>60</b>	<b>36</b>	<b>58</b>	<b>154</b>

## Appendix G – COUNTRY MAP

