

The Republic of Yemen 2010 Enterprise Surveys Data Set

1. Introduction

1. This document provides additional information on the data collected in Yemen between March 2010 and June 2010 as part of the Enterprise Survey, an initiative of the World Bank.

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.

2. Sampling Structure

2. The sample for Yemen was selected using stratified random sampling, following the methodology explained in the *Sampling Manual*¹. Stratified random sampling² was preferred over simple random sampling for several reasons³:

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.)

¹ The complete text can be found at http://www.enterprisesurveys.org/documents/Implementation_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

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 4 manufacturing industries, 1 services industry -retail -, and two residual sectors as defined in the sampling manual. Each manufacturing industry had a target of 55 - 57 interviews. The services industry and the other service residual sectors had a target of 125-135 interviews. For the manufacturing industries sample sizes were inflated by about 50% 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 six Governorates: Aden, Hudaydah, Ibb, Mekhalla, Sana, and Taiz.

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. 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. One frame was used for Yemen. The sample frame containing contacts used in the Yemen was obtained from a subset of the 2004 Population and Establishments Census from Center of Statistics Organization (CSO). Each database contained the following information:

- Name of the firm
- Location
- Contact details (no telephone information is available)
- ISIC code
- Number of employees.

Counts from sample frames shown below.

Universe Figures for Yemen from the 2004 Population and Establishments Census
from Center of Statistics Organization (CSO)

Region	Size	15	18	26	28	Rest of manuf.	Retail	Other services	Grand Total
Aden									
	small(5-19)	115	68	82	48	60	153	696	1,222
	medium(20-99)	3		5	5	12	10	70	105
	large(100 and over)	4			1	10		18	33
Total		122	68	87	54	82	163	784	1,360
Hudaydah									
	small(5-19)	255	70	156	77	84	193	859	1,694
	medium(20-99)	8	7	3	1	9	11	63	102
	large(100 and over)	5			1	8	2	4	20
Total		268	77	159	79	101	206	926	1,816
Ibb									
	small(5-19)	55	14	59	56	20	98	338	640
	medium(20-99)	1		3		1	3	14	22
	large(100 and over)								0
Total		56	14	62	56	21	101	352	662
Mekhalha									
	small(5-19)	66	20	99	68	101	79	535	968
	medium(20-99)	9		11		11	6	36	73
	large(100 and over)	4		1		3		5	13
Total		79	20	111	68	115	85	576	1,054
Sana									
	small(5-19)	313	115	424	320	340	710	2477	4,699
	medium(20-99)	21	7	21	11	31	32	214	337
	large(100 and over)	4	1		1	7	1	24	38
Total		338	123	445	332	378	743	2715	5,074
Taiz									
	small(5-19)	154	152	88	82	58	203	779	1,516
	medium(20-99)	5	4	2	6	12	7	39	75
	large(100 and over)	5		1	1	14		4	25
Total		164	156	91	89	84	210	822	1,616
Grand Total		1027	458	955	678	781	1508	6175	11582

9. The enumerated establishments were then used as the frame for the selection of a sample with the aim of obtaining interviews at 540 establishments with five or more employees

10. 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. In addition, the sample frame contains no telephone/fax numbers so the local contractor had to screen the contacts by visiting them. 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 43% (472 out of 1085 establishments). Breaking down by stratified industries, the following numbers of establishments were surveyed:

15 - Food products and beverages	54
18 - Apparel	45
26 - Non-metallic mineral products	47
28 - Fabricated metal products	46
Other Manufacturing	59
Other Services	140
Retail	86

Local Agency team involved in the study:

Local Agency	Name: Prodigy Systems Location: Sana'a, Yemen Membership of international organization: N/A Activities since: N/A
Name of Project Manager	Mr. Adnan A. Al Harazi
Name and position of other key persons of the project: Local Survey Implementation Team and corresponding supervisor and enumerator codes:	Ms. Safa Almoayad Yemen Country Project Coordinator Others: - Yemen Fieldwork Manager - Lead Supervisor - Data entry Specialist
Other staff involved:	20 Enumerators

Sample Frame:

Characteristics of sample frame used	Variables: Name of establishment, address, activity, number of employees
Source:	Center of Statistics Organization (CSO)
Year:	2004
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. Telephone number is not available. There are some firms missing contact information.
Year and organism who conducted the last economic census	Center of Statistics Organization (CSO)

Other sources for companies statistics	Oral/face-to-face interview
--	-----------------------------

Sectors included in the sample:

Original Sectors	Manufacturing: 15,16,17,19,20,21,22,24,25,26, 28,29,36,37 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: 1 = Sana'a 2 = Aden 3 = Taiz 4 = Al Hudaida 5 = Mukala 6 = Ibb
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	Mar 2010 - June 2010
Locations	Six Governorates: Aden, Hudaydah, Ibb, Mekhalla, Sana, and Taiz.
Interview number	Manufacturing: 244 Services: 91 Other Services: 142
Problems found during fieldwork	Screening is very difficult in Yemen due to: - The sample provided did not provide phone numbers. - Many establishments had no specific name but written by activity like a "Bakery", "Textile Shop", "Grocery", etc. - Many small establishments have no specific contact numbers. - The addresses provided in the sample were vaguely

	<p>described so a screening was important to locate the targeted points to facilitate the job in the fieldwork.</p> <p>Despite the efforts in explaining the neutrality of the survey, establishments were hesitant to give the real number of their staff afraid of leaking the information to the Tax Authority. Some individuals were unwilling to provide income, sales, and expenditure related information.</p>
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 “YE” indicate questions specific to Yemen 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.

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 (*11*, *16* and *18*) 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 *11*, *16* and *18* 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 Yemen 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 *wstrict*.

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

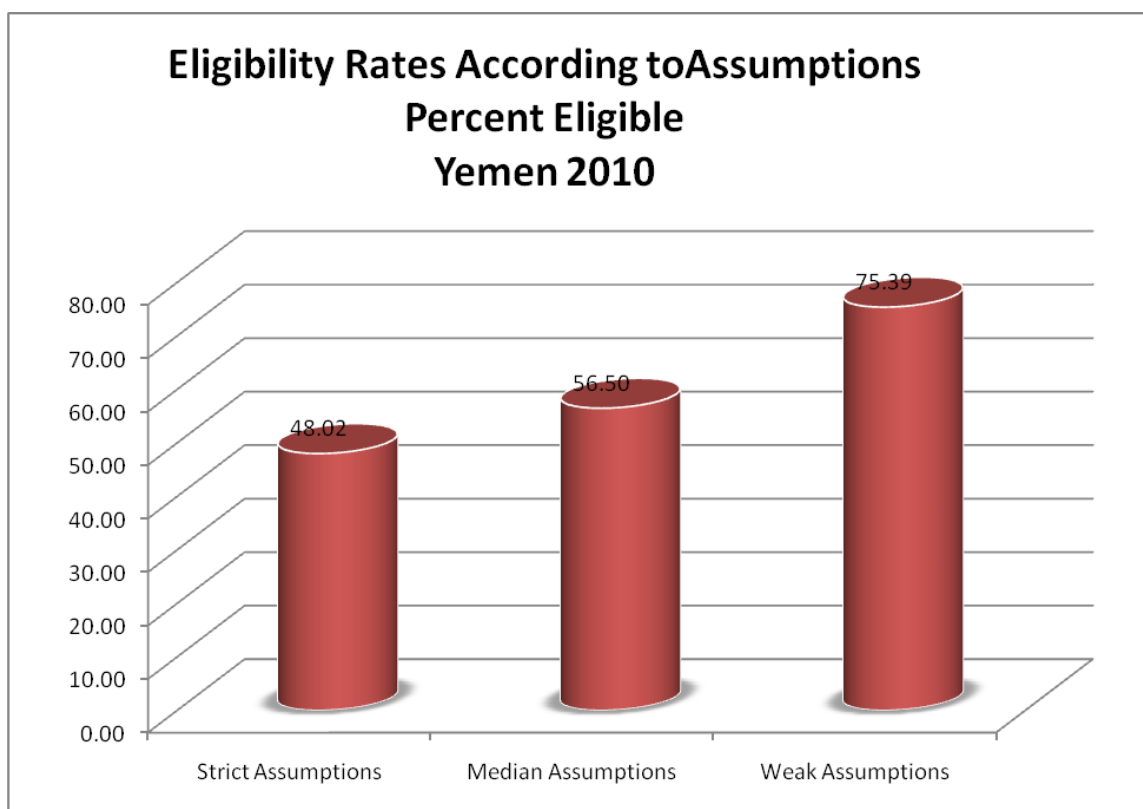
26. Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

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 *wweak*.

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 Yemen 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.)⁴

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

⁴ This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

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 Yemen.

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.)⁵

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.⁶ 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.

8. Non-response

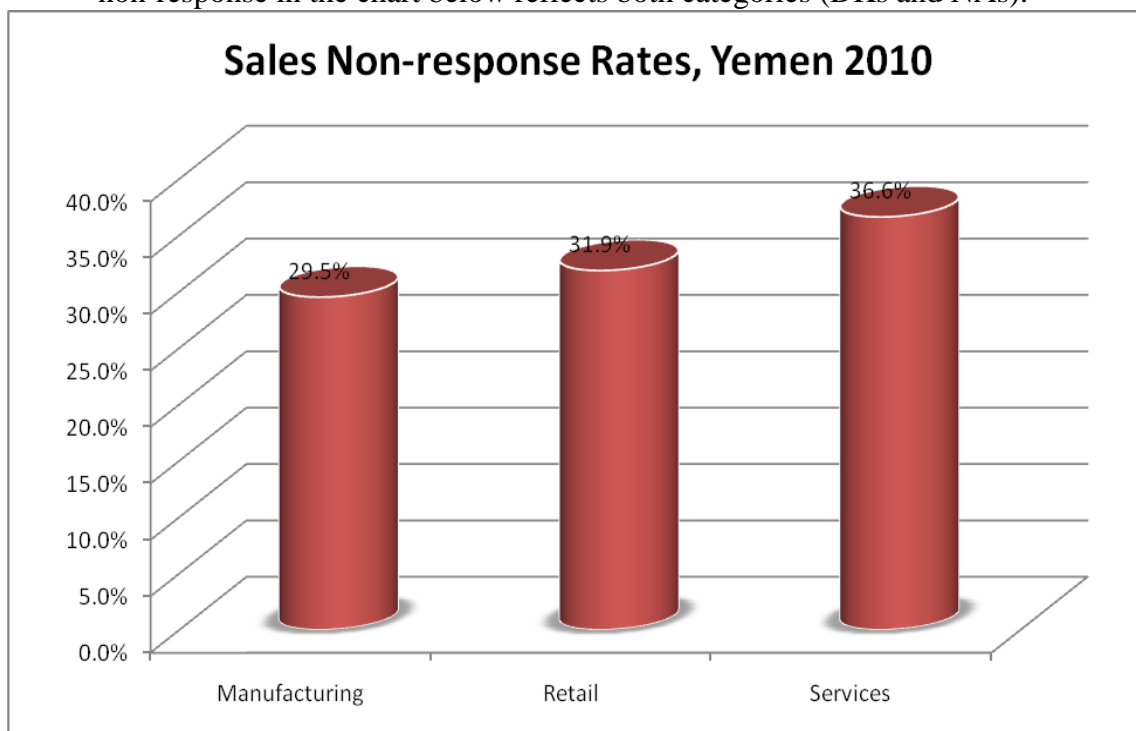
37. Survey non-response must be differentiated from item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the

⁵ 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.

⁶ 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.

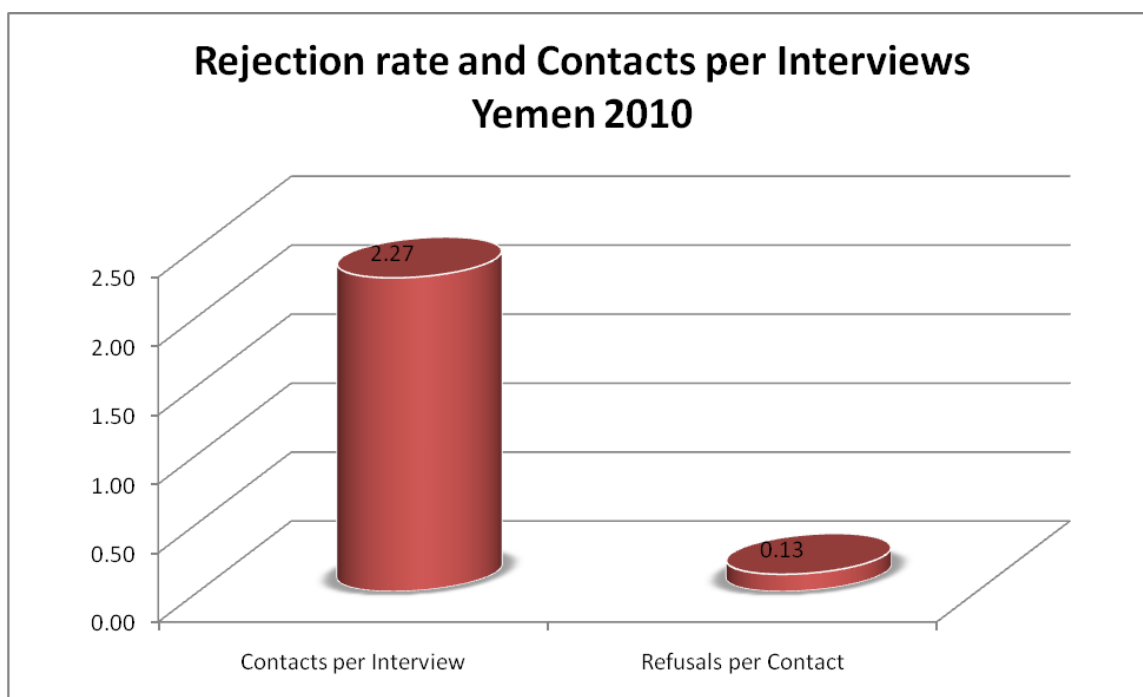
refusals to answer some specific questions. Enterprise Surveys suffer from both problems and different strategies were used to address these issues.

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 2.27. 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 0.13.



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 Yemen. 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)	No problems
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	Prodigy System's mobile application
Comments on the data cleaning	Checking for data errors and inconsistencies was conducted by Prodigy System 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:	<p>a. Yemen has a high rate of illiteracy. Many of the business owners interviewed had modest education that made it difficult for them to understand many concepts in the questionnaire.</p> <p>b. Some small and even medium establishments practice no professional accounting methods nor have any financial bookkeeping. They count their profits day by day. It was difficult to get any accurate productivity data from them</p> <p>c. The taxes system in the country is one of the major economic problems in Yemen for its ineffectiveness. Most of the businesses state wrong figures to tax inspectors. Many businesses would see it a big risk to give any accurate data to anyone else.</p> <p>d. There was a major difficulty in estimating effort and time</p>
---	--

	<p>for field work as people tended to give fake appointments avoiding the embarrassment of saying a firm 'no'. Those made the team go back and forth uselessly to many establishments.</p> <p>e. There is usually a weak and unorganized chain of communication in the structure of establishments in Yemen. Therefore, it is of the least effective mechanism to try to communicate with the managers through phones or faxes. That was proven in the callbacks of re-approaching the refusals and missing data in the final phase as we needed to resend faxes more than one time, getting in a loop of numbers and people with likely no clear response.</p>
Relevant local events occurred during fieldwork:	<ul style="list-style-type: none"> - The fieldwork coincided with the rounds of tax inspection and social security. - Security instability in Aden and Hadramout was a threat. However, field work went smoothly in Aden but was not easy in Hadramout because it coincided with the visit of the President to the city where separatists went out in demonstrations and security was taking strictest measures.
Other aspects:	

Appendix B

Status Codes Fresh:

	Eligibles	
Eligibles	1. Eligible establishment (Correct name and address)	501
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	9
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	4
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	7
Ineligibles	5. The establishment has less than 5 permanent full time employees	83
	6. The firm discontinued businesses	73
	7. Not a business: private household	2
	8. Ineligible activity: education, agriculture, finances, governments...	85
Unobtainable	91. No reply (<i>after having called in different days of the week and in different business hours</i>)	0
	92. Line out of order	0
	93. No tone	0
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	205
	13. Refuses to answer the screener	92
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	0
	151. Out of target - outside the covered regions, firm moved abroad	24
	152. Out of target - firm moved abroad	0
	Total	1085

Response Outcomes Fresh:

Sample Target	540
Complete interviews (Total)	477
Incomplete interviews	0
Eligible in process	0
Refusals	44
Out of target	243
Impossible to contact	205
Ineligible - coop.	24
Refusal to the Screener	92
Total Used	1085

Appendix C

Universe Estimates, Yemen:

Region	Size	15	18	26	28	Rest of manuf.	Retail	Other services	Grand Total
Aden									
	small(5-19)	115	68	82	48	60	153	696	1,222
	medium(20-99)	3		5	5	12	10	70	105
	large(100 and over)	4			1	10		18	33
Total		122	68	87	54	82	163	784	1,360
Hudaydah									
	small(5-19)	255	70	156	77	84	193	859	1,694
	medium(20-99)	8	7	3	1	9	11	63	102
	large(100 and over)	5			1	8	2	4	20
Total		268	77	159	79	101	206	926	1,816
Ibb									
	small(5-19)	55	14	59	56	20	98	338	640
	medium(20-99)	1		3		1	3	14	22
	large(100 and over)								0
Total		56	14	62	56	21	101	352	662
Mekhallah									
	small(5-19)	66	20	99	68	101	79	535	968
	medium(20-99)	9		11		11	6	36	73
	large(100 and over)	4		1		3		5	13
Total		79	20	111	68	115	85	576	1,054
Sana									
	small(5-19)	313	115	424	320	340	710	2477	4,699
	medium(20-99)	21	7	21	11	31	32	214	337
	large(100 and over)	4	1		1	7	1	24	38
Total		338	123	445	332	378	743	2715	5,074
Taiz									
	small(5-19)	154	152	88	82	58	203	779	1,516
	medium(20-99)	5	4	2	6	12	7	39	75
	large(100 and over)	5		1	1	14		4	25
Total		164	156	91	89	84	210	822	1,616
Grand Total		1027	458	955	678	781	1508	6175	11582

Appendix D

Strict Cell Weights Yemen:

Region	Size	15 Food	18 Apparel	26 non-metallic mineral	28 fabricated metal products	Rest of manuf.	Retail	Other services
Aden								
	small(5-19)	18.71	6.39	18.53	10.93	49.35	5.82	125.99
	medium(20-99)	2.39		1.23	1.86	1.61	1.96	2.87
	large(100 and over)	0.52			0.37	1.66		2.19
Total								
Hudaydah								
	small(5-19)	29.92	4.43	14.83	6.56	11.63	7.82	130.84
	medium(20-99)	2.68	0.90	0.93	0.62	2.54	1.36	3.01
	large(100 and over)				0.27	0.98	0.43	
Total								
Ibb								
	small(5-19)	10.02	1.47	9.95	5.71	3.07	5.66	28.54
	medium(20-99)			0.98				1.31
	large(100 and over)							
Total								
Mekhalha								
	small(5-19)		3.37	20.05	20.81	55.83	4.26	162.71
	medium(20-99)	1.35		1.91		1.74	1.66	3.58
	large(100 and over)	3.50				1.11		
Total								
Sana								
	small(5-19)		5.29		9.73	51.35	17.13	41.17
	medium(20-99)	1.40	0.84	1.01	0.60	1.75	0.79	6.63
	large(100 and over)	1.54				0.57		0.61
Total								
Taiz								
	small(5-19)	6.87	7.83	5.45	6.14	19.61	4.31	36.23
	medium(20-99)	0.89	0.82	0.49	1.00	1.35		1.61
	large(100 and over)	1.17				1.83		0.86

Median Cell Weights Yemen:

Region	Size	15 Food	18 Apparel	26 non-metallic mineral	fabricated metal products	Rest of manuf.	Retail	Other services
Aden								
	small(5-19)	17.88	5.00	16.54	9.19	48.85	5.18	125.37
	medium(20-99)	2.41		1.16	1.65	1.68	1.84	3.02
	large(100 and over)	0.49			0.30	1.59		2.11
Total								
Hudaydah								
	small(5-19)	41.26	5.00	19.11	7.96	16.61	10.05	187.92
	medium(20-99)	4.77	1.32	1.55	0.98	4.68	2.26	5.58
	large(100 and over)				1.31	5.58	2.20	
Total								
Ibb								
	small(5-19)	11.39	1.37	10.57	5.71	3.62	5.99	33.79
	medium(20-99)			1.13				1.69
	large(100 and over)							
Total								
Mekhalha								
	small(5-19)		3.29	22.33	21.84	68.96	4.73	202.06
	medium(20-99)	2.03		2.68		2.70	2.33	5.60
	large(100 and over)	4.78				1.56		
Total								
Sana								
	small(5-19)		4.59		9.07	56.38	16.90	45.44
	medium(20-99)	1.53	0.75	1.03	0.58	1.98	0.81	7.55
	large(100 and over)	2.28				0.87		0.95
Total								
Taiz								
	small(5-19)	7.62	7.12	5.65	6.00	22.56	4.45	41.90
	medium(20-99)	1.10	0.83	0.57	1.08	1.73		2.07
	large(100 and over)	1.58				2.58		1.22

Weak Cell Weights Yemen:

Region	Size	15 Food	18 Apparel	26 non-metallic mineral	fabricated metal products	Rest of manuf.	Retail	Other services
Aden								
	small(5-19)	14.37	5.68	15.85	8.86	39.81	4.93	94.65
	medium(20-99)	1.99		1.14	1.63	1.41	1.81	2.34
	large(100 and over)	0.36			0.27	1.20		1.47
Total								
Hudaydah								
	small(5-19)	38.04	6.51	20.99	8.80	15.52	10.97	162.66
	medium(20-99)	4.28	1.67	1.66	1.05	4.26	2.40	4.71
	large(100 and over)				1.23	4.44	2.05	
Total								
Ibb								
	small(5-19)	15.67	2.66	17.33	9.43	5.04	9.77	43.66
	medium(20-99)			1.26				1.48
	large(100 and over)							
Total								
Mekhalha								
	small(5-19)		3.57	20.42	20.10	53.65	4.30	145.61
	medium(20-99)	1.44		2.26		1.94	1.96	3.72
	large(100 and over)	3.39				1.12		
Total								
Sana								
	small(5-19)		17.46		29.28	153.78	53.89	114.81
	medium(20-99)	3.05	2.11	2.45	1.38	3.99	1.90	14.10
	large(100 and over)	3.10				1.20		1.21
Total								
Taiz								
	small(5-19)	11.13	14.67	9.83	10.50	33.37	7.70	57.41
	medium(20-99)	1.18	1.26	0.73	1.39	1.87		2.08
	large(100 and over)	1.47				2.43		1.07

Appendix E

Original Sample Design, Yemen:

Region	Size	15	18	26	28	Rest of manuf.	Retail	Other services	Grand Total
Aden									
	small(5-19)	4	8	5	3	1	16	5	42
	medium(20-99)	2	0	3	3	4	5	10	27
	large(100 and over)	4	0	0	1	5	0	8	18
Total		10	8	8	7	10	21	23	87
Hudaydah									
	small(5-19)	4	7	6	5	3	17	6	48
	medium(20-99)	3	4	4	1	4	6	10	32
	large(100 and over)	4	0	0	1	3	2	4	14
Total		11	11	10	7	10	25	20	94
Ibb									
	small(5-19)	3	4	3	5	3	8	5	31
	medium(20-99)	0	0	2	0	0	2	5	9
	large(100 and over)	0	0	0	0	0	0	0	0
Total		3	4	5	5	3	10	10	40
Mekhalha									
	small(5-19)	0	4	6	5	3	10	5	33
	medium(20-99)	6	0	6	0	5	6	10	33
	large(100 and over)	3	0	0	0	2	0	6	11
Total		9	4	12	5	10	16	21	77
Sana									
	small(5-19)	0	10	4	13	2	15	12	56
	medium(20-99)	8	5	8	7	7	15	12	62
	large(100 and over)	4	1	0	1	5	2	15	28
Total		12	16	12	21	14	32	39	146
Taiz									
	small(5-19)	7	10	6	6	3	19	6	57
	medium(20-99)	2	2	2	4	4	2	10	26
	large(100 and over)	2	0	2	0	3	0	6	13
Total		11	12	10	10	10	21	22	96
Grand Total		56	55	57	55	57	125	135	540

Completed Interviews, Yemen:

Region	Size	15	18	26	28	Rest of m	Retail	Other ser	Grand Total
Aden									
	small(5-19)	6	8	4	4	1	19	5	47
	medium(20-99)	1		3	2	5	3	18	32
	large(100 and over)	3			1	2		3	9
Total		10	8	7	7	8	22	26	88
Hudaydah									
	small(5-19)	7	10	8	9	5	15	5	59
	medium(20-99)	2	4	2	1	2	4	13	28
	large(100 and over)				1	2	1		4
Total		9	14	10	11	9	20	18	91
Ibb									
	small(5-19)	3	4	3	5	3	7	6	31
	medium(20-99)			2				7	9
	large(100 and over)								0
Total		3	4	5	5	3	7	13	40
Mekhalha									
	small(5-19)		3	3	2	1	9	2	20
	medium(20-99)	5		4		4	2	7	22
	large(100 and over)	1				2			3
Total		6	3	7	2	7	11	9	45
Sana									
	small(5-19)		6		11	2	11	20	50
	medium(20-99)	7	3	9	8	7	14	14	62
	large(100 and over)	1				4		14	19
Total		8	9	9	19	13	25	48	131
Taiz									
	small(5-19)	9	6	6	5	1	14	8	49
	medium(20-99)	3	2	2	3	4		12	26
	large(100 and over)	2				3		2	7
Total		14	8	8	8	8	14	22	82
Grand Total		50	46	46	52	48	99	136	477