

Sri Lanka 2011 Enterprise Surveys Data Set

1. Introduction

1. This document provides additional information on the data collected in Sri Lanka between June 2011 and November 2011 as part of the Enterprise Survey component of two surveys carried out in Sri Lanka simultaneously in 2011.

The objective of the survey is to obtain feedback from enterprises on the state of the private sector in client countries 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 Sri Lanka was selected using stratified random sampling, following the methodology explained in the Sampling Manual¹. 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.)

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.

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

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 2 manufacturing industries, 1 services industry -retail -, and two residual sectors as defined in the sampling manual. Each manufacturing industry had a target of 120 interviews. The services industry and the two residual sectors had a target of 120 interviews.

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 nine regions: Eastern, Western, Southern, Central, Northern, North-Central, North-West, Uva, Sabaragamuwa.

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. One frame was used for the Enterprise Survey in Sri Lanka. The sample frame containing fresh contacts used in the Sri Lanka was obtained from the Department of Census and Statistics of Sri Lanka (DCS) 2003. The database contained the following information:

- Name of the firm
- Location
- Contact details
- ISIC code
- Number of employees.

Counts from sample frames are shown below.

Estimated Universe for Sri Lanka Enterprise Survey

ESTIMATED UNIVERSE							
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail	Grand Total
Central	1-Small	425	110	1039	1305	981	3860
	2-Medium	214	18	98	151	98	579
	3-Large	45	44	28	48	3	168
	TOTAL	684	172	1165	1504	1082	4607
Eastern	1-Small	349	58	345	386	348	1486
	2-Medium	20	2	7	38	32	99
	3-Large	1	3	5	14	3	26
	TOTAL	370	63	357	438	383	1611
North-Central	1-Small	413	38	150	516	342	1459
	2-Medium	30	2	17	50	25	124
	3-Large	4	18	5	10	1	38
	TOTAL	447	58	172	576	368	1621
Northern	1-Small	302	49	173	371	382	1277
	2-Medium	32	0	15	30	32	109
	3-Large	0	0	2	8	2	12
	TOTAL	334	49	190	409	416	1398
North-West	1-Small	608	103	2411	980	610	4712
	2-Medium	77	13	161	82	40	373
	3-Large	13	48	32	24	5	122
	TOTAL	698	164	2604	1086	655	5207
Sabaragamuwa	1-Small	170	71	475	617	396	1729
	2-Medium	109	14	79	47	37	286
	3-Large	47	23	17	14	1	102
	TOTAL	326	108	571	678	434	2117
Southern	1-Small	295	123	881	942	598	2839
	2-Medium	185	15	70	102	72	444
	3-Large	46	45	21	29	5	146
	TOTAL	526	183	972	1073	675	3429
Uva	1-Small	99	20	185	448	282	1034
	2-Medium	68	4	9	46	19	146
	3-Large	9	8	3	10	1	31
	TOTAL	176	32	197	504	302	1211
Western	1-Small	1030	732	3667	5066	3869	14364
	2-Medium	192	193	863	903	559	2710
	3-Large	87	213	333	173	56	862
	TOTAL	1309	1138	4863	6142	4484	17936
Grand Total		4870	1967	11091	12410	8799	39137

9. The enumerated establishments were then used as the frame for the selection of a sample with the aim of obtaining interviews at 600 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 DCS and 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 50% (900 out of 1806 establishments).⁴ Breaking down by industry, the following numbers of establishments were surveyed:

15 (Food)	121
18 (Garments)	120
Other manufacturing	122
Retail	120
Other services	127

Local Agency team involved in the study:

Local Agency	Name: The Nielsen Company Lanak (Pvt) Ltd Location: Colombo, Sri Lanka
Name of Project Manager	Mr. Jinendra Kothalawala
Other staff involved:	100+ enumerators and screeners

Sample Frame:

Characteristics of sample frame used	Variables: Name of establishment, address, activity, telephone number, number of employees
Source:	Sri Lanka Department of Census and Statistics
Year:	2003
Comments on the quality of sample frame:	Contained all the variables needed to construct a sample frame for the enterprise survey. Several addresses and telephone number outdated. There are some firms missing contact information.

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

Sectors included in the sample:

Original Sectors	Food & Beverage Manufacturing (ISIC Code 15)
	Garment Manufacturing (ISIC Code 18)
	Other Manufacturing (ISIC Code 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 33, 34, 35, 36, 37)
	Retail (ISIC Code 52)
	Other Services (ISIC Code 50, 51, 55, 60, 61, 62, 63, 64, 72)

Sample:

Comments/ problems on sectors and regions selected in the sample	DCS Sampling frames namely, Census of Industry –listing was 2003 and Trade and service was 2003/6 and some of those were not updated. Hence closedown business, business place changes etc. were observed. Since telephone numbers were not available in some businesses and only postal addresses were available tracing the place was difficult for some businesses.
Comments on the response rate	
Comments on the sample design:	
Other comments:	None

Fieldwork:

Date of Fieldwork	June 8 th , 2011-November 2 nd , 2011
Interview number	Manufacturing: 363 Services: 120 Other Services: 127
Problems found during fieldwork	Due to tax, labor regulations and other issues some respondents were reluctant to reveal the actual figures as observed by ourselves in similar surveys.
Other observations:	None

4. Data Base Structure:

12. The structure of the data base reflects the fact that 2 different versions of the questionnaire were used. The first variation, the Manufacturing Questionnaire, is built upon a core set of questions and adds some specific questions relevant to the sector. The second expanded variation, the Services Questionnaire, also 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 “SA” indicate questions specific to South Asia 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 *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 Sri Lanka 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

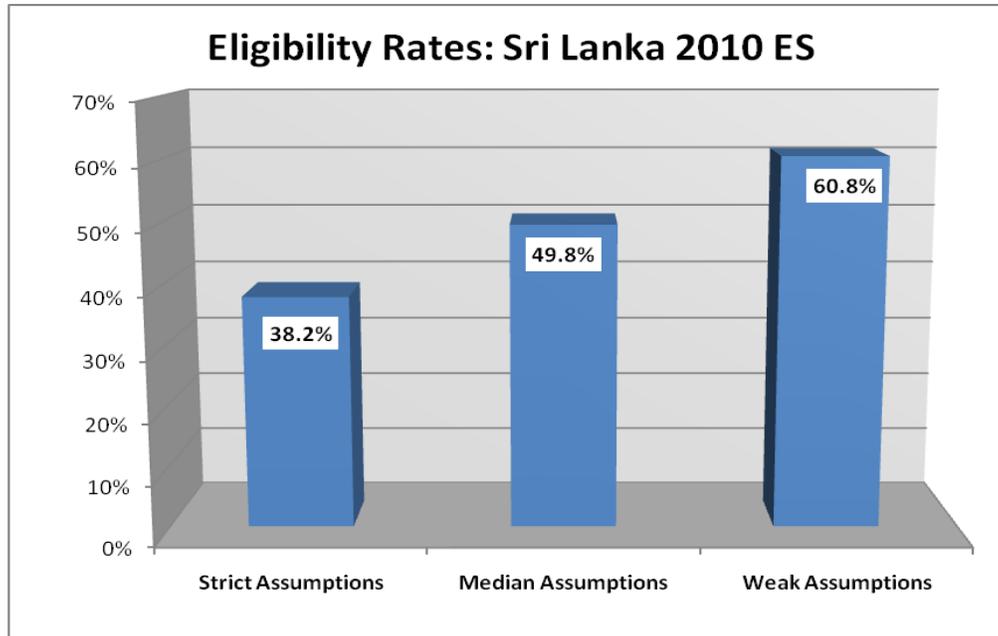
26. Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *w_median*.

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, 94,10,11, 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 Sri Lanka 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.)⁵

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

⁵ 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 Sri Lanka.

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.)⁶

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.⁷ 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

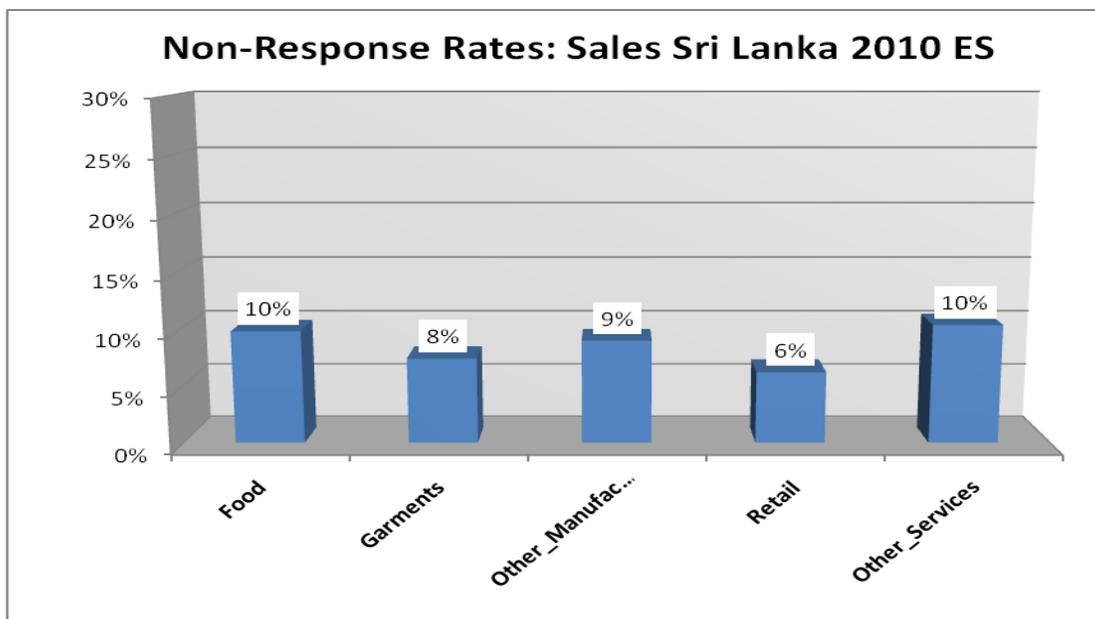
37. Survey non-response must be differentiated from item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the

⁶ 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.

⁷ 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.

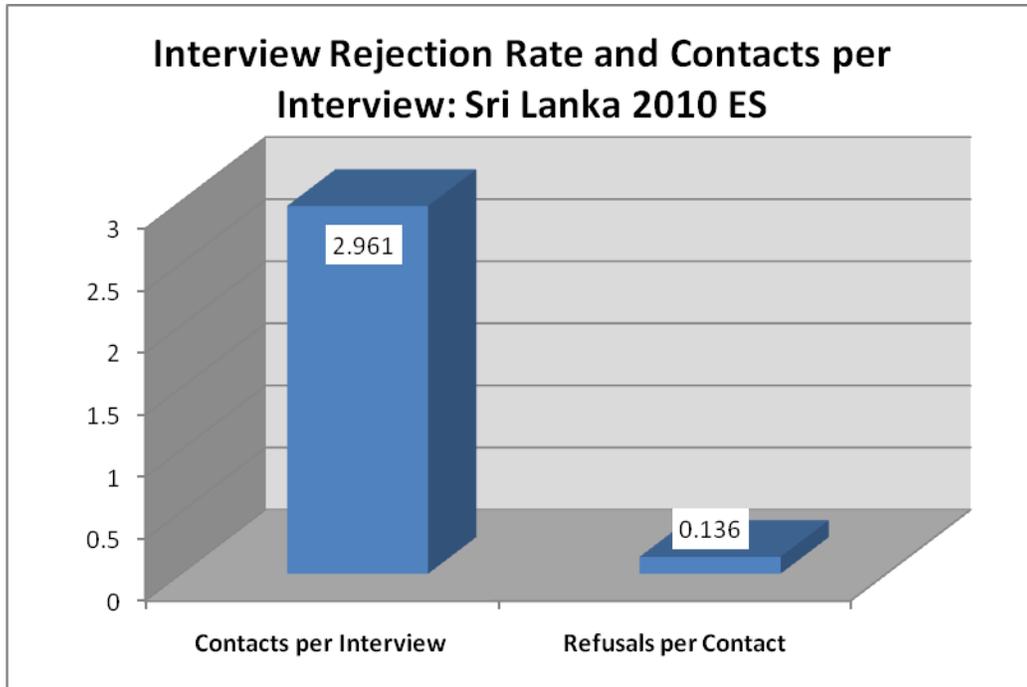
refusals to answer some specific questions. Enterprise Surveys suffer from both problems and different strategies were used to address these issues.

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 2.96. 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 0.136.



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 Sri Lanka. 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:	No problems
Suggestions or other comments on the questionnaire:	No problems

Database:

Comments on the data entry program	Data entry program chosen: CSPro
Comments on the data cleaning	Checking for data errors and inconsistencies was conducted by Nielsen and the World Bank. 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:	
Relevant local events occurred during fieldwork:	

Other aspects:	
----------------	--

Appendix B

Status Codes:

Sri Lanka		
Eligibles	1. Eligible establishment (Correct name and address)	666
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	11
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	6
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	6
Ineligibles	5. The establishment has less than 5 permanent full time employees	77
	6. The firm discontinued businesses	165
	7. Not a business: private household	5
	8. Ineligible activity: education, agriculture, finances, governments...	202
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	16
	92. Line out of order	5
	93. No tone	1
	94. Phone number does not exist	4
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	172
	13. Refuses to answer the screener	211
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	246
	16. duplicate entry	24
	151. Out of target - outside the covered regions, firm moved abroad	7
	152. Out of target - firm moved abroad	0
	153. Out of target - Not registered with SAT	6
	Total	1830

Response Outcomes:

	Sri Lanka
Sample Target	600
Complete interviews (Total)	610
Incomplete interviews	3
Elegible in process	33
Refusals	43
Out of target	449
Impossible to contact	198
Ineligible - coop.	13
Refusal to the Screener	211
Total	1560

Appendix C

Universe Estimates, Sri Lanka:

ESTIMATED UNIVERSE							
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail	Grand Total
Central	1-Small	425	110	1039	1305	981	3860
	2-Medium	214	18	98	151	98	579
	3-Large	45	44	28	48	3	168
	TOTAL	684	172	1165	1504	1082	4607
Eastern	1-Small	349	58	345	386	348	1486
	2-Medium	20	2	7	38	32	99
	3-Large	1	3	5	14	3	26
	TOTAL	370	63	357	438	383	1611
North-Central	1-Small	413	38	150	516	342	1459
	2-Medium	30	2	17	50	25	124
	3-Large	4	18	5	10	1	38
	TOTAL	447	58	172	576	368	1621
Northern	1-Small	302	49	173	371	382	1277
	2-Medium	32	0	15	30	32	109
	3-Large	0	0	2	8	2	12
	TOTAL	334	49	190	409	416	1398
North-West	1-Small	608	103	2411	980	610	4712
	2-Medium	77	13	161	82	40	373
	3-Large	13	48	32	24	5	122
	TOTAL	698	164	2604	1086	655	5207
Sabaragamuwa	1-Small	170	71	475	617	396	1729
	2-Medium	109	14	79	47	37	286
	3-Large	47	23	17	14	1	102
	TOTAL	326	108	571	678	434	2117
Southern	1-Small	295	123	881	942	598	2839
	2-Medium	185	15	70	102	72	444
	3-Large	46	45	21	29	5	146
	TOTAL	526	183	972	1073	675	3429
Uva	1-Small	99	20	185	448	282	1034
	2-Medium	68	4	9	46	19	146
	3-Large	9	8	3	10	1	31
	TOTAL	176	32	197	504	302	1211
Western	1-Small	1030	732	3667	5066	3869	14364
	2-Medium	192	193	863	903	559	2710
	3-Large	87	213	333	173	56	862
	TOTAL	1309	1138	4863	6142	4484	17936
Grand Total		4870	1967	11091	12410	8799	39137

Appendix D

Strict Cell Weights Sri Lanka:

STRICT Weights						
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail
Central	1-Small	55.3		77.1	62.2	47.8
	2-Medium	30.2		15.8	8.9	4.7
	3-Large	3.1		11.5	4.2	
Eastern	1-Small	24.0		20.3	15.6	89.7
	2-Medium	1.2		0.9	4.2	1.8
	3-Large					
North-Central	1-Small	47.4		19.7	86.9	24.5
	2-Medium	2.5		1.8	6.1	1.9
	3-Large				1.6	
Northern	1-Small	24.2		5.9	16.4	43.1
	2-Medium	2.8		1.7	1.9	2.6
	3-Large					
North-West	1-Small	36.3	6.7	123.3	64.4	41.0
	2-Medium	5.4	1.3	8.9	7.8	3.6
	3-Large	2.2	10.5	2.6	5.8	
Sabaragamuwa	1-Small	17.1		54.5	36.4	49.7
	2-Medium	23.7		9.8	5.0	3.8
	3-Large	4.4		1.4		
Southern	1-Small	30.7		117.7	86.2	31.8
	2-Medium	31.3		6.8	7.2	6.5
	3-Large	6.6		3.5	4.1	
Uva	1-Small	12.2		14.7	30.5	19.6
	2-Medium	13.7		0.8	3.4	2.4
	3-Large					
Western	1-Small	51.7	5.8	112.5	186.4	182.0
	2-Medium	11.6	4.2	36.6	36.0	21.9
	3-Large	6.7	9.8	33.1	4.5	2.7

Median Cell Weights Sri Lanka:

MEDIAN Weights						
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail
Central	1-Small	76.6		108.3	78.0	58.4
	2-Medium	48.6		25.7	13.0	6.7
	3-Large	5.9		22.3	7.3	
Eastern	1-Small	24.4		20.9	14.3	80.4
	2-Medium	1.4		1.1	4.4	1.9
	3-Large					
North-Central	1-Small	56.5		23.7	93.6	25.8
	2-Medium	3.4		2.5	7.6	2.4
	3-Large				2.3	
Northern	1-Small	30.4		7.5	18.6	47.6
	2-Medium	4.1		2.5	2.5	3.3
	3-Large					
North-West	1-Small	35.5	5.8	122.2	57.0	35.3
	2-Medium	6.2	1.3	10.3	8.0	3.6
	3-Large	2.9	12.6	3.5	7.1	
Sabaragamuwa	1-Small	27.0		87.2	52.0	69.3
	2-Medium	43.6		18.3	8.3	6.1
	3-Large	9.5		3.0		
Southern	1-Small	31.0		120.5	78.8	28.3
	2-Medium	36.8		8.0	7.7	6.8
	3-Large	9.2		4.9	5.2	
Uva	1-Small	17.6		21.4	39.7	24.9
	2-Medium	22.9		1.3	5.1	3.5
	3-Large					
Western	1-Small	69.8	7.0	154.0	227.8	216.6
	2-Medium	18.2	5.8	58.1	51.2	30.3
	3-Large	12.5	16.2	62.3	7.6	4.4

Weak Cell Weights Sri Lanka:

WEAK Weights						
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail
Central	1-Small	86.2		131.8	94.5	67.9
	2-Medium	46.6		26.7	13.4	6.6
	3-Large	5.4		22.2	7.2	
Eastern	1-Small	59.5		55.2	37.6	202.4
	2-Medium	2.9		2.4	9.9	4.0
	3-Large					
North-Central	1-Small	65.4		29.7	116.7	30.8
	2-Medium	3.4		2.7	8.1	2.4
	3-Large				2.4	
Northern	1-Small	85.3		22.9	56.1	138.1
	2-Medium	9.7		6.4	6.5	8.3
	3-Large					
North-West	1-Small	41.7	7.0	155.1	72.0	42.8
	2-Medium	6.2	1.3	11.1	8.6	3.8
	3-Large	2.8	12.4	3.7	7.3	
Sabaragamuwa	1-Small	31.9		111.5	66.1	84.5
	2-Medium	43.9		19.9	9.0	6.4
	3-Large	9.2		3.1		
Southern	1-Small	35.3		148.5	96.6	33.3
	2-Medium	35.7		8.4	8.0	6.8
	3-Large	8.6		4.9	5.2	
Uva	1-Small	20.4		26.8	49.4	29.7
	2-Medium	22.6		1.4	5.4	3.6
	3-Large					
Western	1-Small	85.8	8.8	204.7	301.3	275.0
	2-Medium	19.1	6.3	65.9	57.7	32.8
	3-Large	12.6	16.7	67.7	8.2	4.6

Appendix E

Original Sample Design, Sri Lanka:

The original aim was to obtain 600 interviews consisting of; 120 interviews within each of the sectors (15-Food, 18-Garments, other manufacturing, retail establishments, and other services).

DESIGN							
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail	Grand Total
Central	1-Small	4	0	4	5	9	22
	2-Medium	4	0	4	6	9	23
	3-Large	4	0	4	5	0	13
	Region Total	12	0	12	16	18	58
Eastern	1-Small	4	0	5	4	4	16
	2-Medium	4	0	3	4	4	16
	3-Large	0	0	1	2	0	2
	Region Total	8	0	9	10	8	34
North-Central	1-Small	4	0	4	4	4	16
	2-Medium	4	0	4	4	4	16
	3-Large	0	0	0	1	0	1
	Region Total	8	0	8	9	8	33
Northern	1-Small	6	0	7	4	4	20
	2-Medium	6	0	5	3	4	19
	3-Large	0	0	0	1	0	1
	Region Total	12	0	12	8	8	40
North-West	1-Small	6	12	6	5	8	37
	2-Medium	6	6	6	5	8	31
	3-Large	6	2	6	5	0	19
	Region Total	18	20	18	15	16	87
Sabaragamuwa	1-Small	4	0	4	4	4	16
	2-Medium	4	0	4	4	4	16
	3-Large	4	0	4	2	0	10
	Region Total	12	0	12	10	8	42
Southern	1-Small	4	0	4	5	8	21
	2-Medium	4	0	4	5	8	21
	3-Large	4	0	4	5	0	13
	Region Total	12	0	12	15	16	55
Uva	1-Small	4	0	4	4	4	16
	2-Medium	4	0	3	4	4	16
	3-Large	0	0	1	1	0	1
	Region Total	8	0	8	9	8	33
Western	1-Small	12	60	12	9	10	103
	2-Medium	9	30	9	9	10	67
	3-Large	9	10	9	10	10	48
	Region Total	30	100	30	28	30	218
Grand Total		120	120	120	120	120	600

A total of 1806 eligible establishments were enumerated. The World Bank drew sample and issued replacements. The contacts were issued and given preferences from 1 to 6. 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, Sri Lanka:

Realized Complete Interviews							
Region	Size	Food Manu.	Garment Manu.	Other Manu.	Other Services	Retail	Grand Total
Central	1-Small	4		6	8	10	28
	2-Medium	4		3	7	11	25
	3-Large	7		1	4		12
	Region Total	15		10	19	21	65
Eastern	1-Small	4		4	5	1	14
	2-Medium	5		2	2	5	14
	3-Large						
	Region Total	9		6	7	6	28
North-Central	1-Small	4		3	2	6	15
	2-Medium	6		4	3	6	19
	3-Large				2		2
	Region Total	10		7	7	12	36
Northern	1-Small	3		6	4	2	15
	2-Medium	3		2	3	3	11
	3-Large						
	Region Total	6		8	7	5	26
North-West	1-Small	12	11	12	8	10	53
	2-Medium	11	8	12	6	8	45
	3-Large	4	3	7	2		16
	Region Total	27	22	31	16	18	114
Sabaragamuwa	1-Small	4		3	5	3	15
	2-Medium	2		3	3	4	12
	3-Large	4		4			8
	Region Total	10		10	8	7	35
Southern	1-Small	6		4	5	11	26
	2-Medium	4		6	7	7	24
	3-Large	4		3	3		10
	Region Total	14		13	15	18	60
Uva	1-Small	3		4	4	5	16
	2-Medium	2		4	4	3	13
	3-Large						
	Region Total	5		8	8	8	29
Western	1-Small	10	63	14	10	10	107
	2-Medium	9	25	11	10	13	68
	3-Large	6	10	4	13	9	42
	Region Total	25	98	29	33	32	217
Grand Total		121	120	122	120	127	610