

# The El Salvador 2010 Enterprise Surveys Data Set

## I. Introduction

1. This document provides additional information on the data collected in El Salvador between July 8 2010 and April 19 2011 as part of the Latin America and Caribbean (LAC) Enterprise Survey 2010, 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 methodology, the sample 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 El Salvador was selected using stratified random sampling, following the methodology explained in the *Sampling Note*<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/locations.

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

---

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

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

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

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

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

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

4. Industry stratification was designed in the way that follows: the universe was stratified into 1 manufacturing industry, 1 service industry -retail -, and 1 residual sector as defined in the sampling manual. The manufacturing industry, service industry, and residual sectors had a target each of 120 interviews.

5. Size stratification was defined following the standardized definition for the Enterprise Surveys: 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 two locations (city and the surrounding business area): San Salvador and the Rest of the Country.

### **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 location) 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. TNS Opinion was hired to implement the LAC 2010 enterprise surveys roll out. In El Salvador the local subcontractors were Mercaplan and Data Research, S.A. de C.V.

9. For El Salvador, two sample frames were used. The first was supplied by the World Bank and consists of enterprises interviewed in El Salvador 2006. The World Bank required that attempts should be made to re-interview establishments responding to the El Salvador 2006 survey where they were within the selected geographical locations and met eligibility criteria. That sample is referred to as the Panel. The second sample frame was produced from the Directorio de Empresas de bases de solvencias from 2008 via the Dirección General de Estadísticos y Censos. Each database contained the following information:

- Coverage;
- Up to datedness;
- Availability of detailed stratification variables ;
- Location identifiers- address, phone number, email;

- Electronic format availability;
- Contact name(s).

Counts from sample frames are shown below.

## Panel sample counts

Panel Sample Count (Panel Sample Frame)

Region name	Employees	Manufacturing	Retail	Other Services	Grand Total
Rest of Country	5-19	67	17	16	100
	20-99	48	18	20	86
	100+	45	8	7	60
Rest of Country Total		160	43	43	246
San Salvador	5-19	123	34	28	185
	20-99	107	33	28	168
	100+	46	14	34	94
San Salvador Total		276	81	90	447
Grand Total		436	124	133	693

## Sample Frames

Source: Dirección General de Estadística y Censos, 2008

Region name	Employees	Manufacturing	Retail	Other Services	Grand Total
Rest of Country	5-19	255	847	1062	2164
	20-99	177	152	353	682
	100+	133	23	96	252
Rest of Country Total		565	1022	1511	3098
San Salvador	5-19	403	722	1517	2642
	20-99	267	212	582	1061
	100+	135	32	135	302
San Salvador Total		805	966	2234	4005
Grand Total		1370	1988	3745	7103

10. The two sample frames were then used for the selection of a sample with the aim of obtaining interviews with 360 establishments with five or more employees

11. The quality of the frame was assessed at the outset 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.

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 20.56% (26 out of 917 establishments)<sup>4</sup>.

### 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 (manufacturing, retail, and other services). 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*.

---

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

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 “LAC” indicate questions specific to LAC, 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 2 establishment identifiers, *idstd* and *id*. The first is a global unique identifier. The second is a country unique identifier. The variables *a2* (sampling location), *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 location. Different combinations of these variables generate the strata cells for each industry/location/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 locations

-*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), other manufacturing (2), retail (52), and (45, 50, 51, 55, 60, 63, 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. However, the phone numbers were unavailable in the sample frame, and thus the enumerators applied the screeners in person. The variables *a4b* and *a6b* contain the industry and size of the establishment from the screener questionnaire. Variables *a8* to *a11* contain additional information and were also collected in the screening phase.

19. Note that there are additional variables for location size by population (*a3*) and firm size by number of workers (*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. 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.

21. Variable *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**

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

23. Appendix B shows the overall estimates of the numbers of establishments in El Salvador based on the sample frame.

24. 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.

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

26. Strict assumption: eligible establishments are only those for which it was possible to directly determine eligibility. The resulting weights, which include adjustments applied to panel firms (see below), are included in the variable *w\_strict\_panadj*.

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

27. 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\_panadj*.

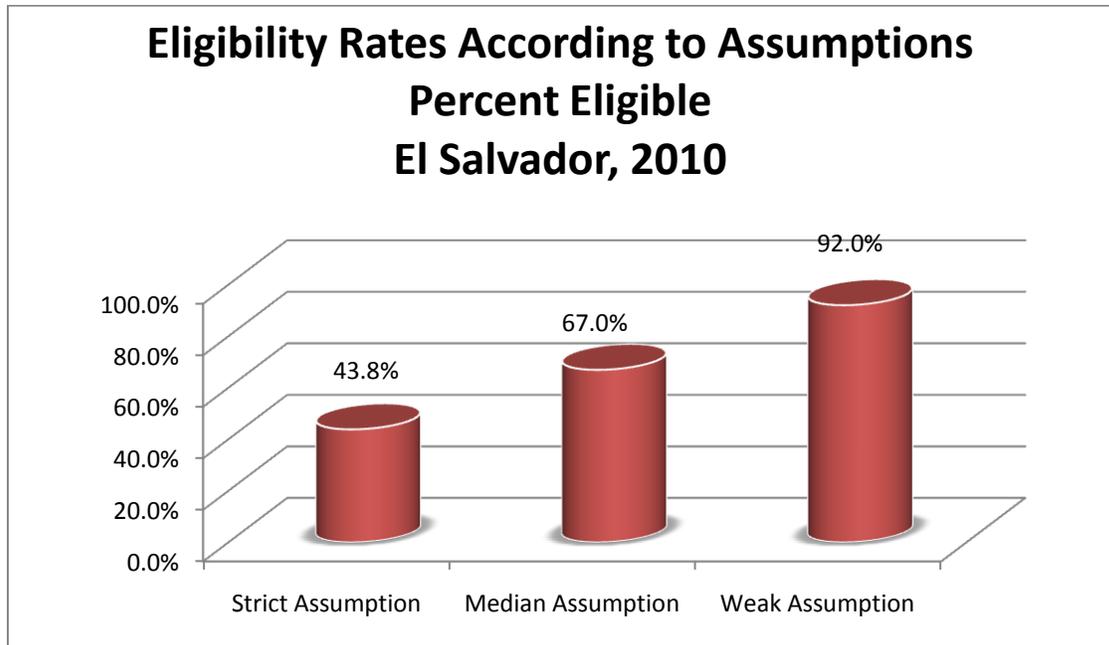
Median eligibility = (Sum of the firms with codes 1,2,3,4,16,10,11, & 13) / Total

28. 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  $w\_weak\_panadj$ .

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

29. 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.



30. Universe estimates for the number of establishments in each industry-location-size cell in El Salvador 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.

31. 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

32. 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>

33. Special care was given to the correct computation of the weights. It was imperative to accurately adjust the totals within each location/industry/size stratum to account for the presence of ineligible units (the firm discontinued business 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 in the phone line, answering machine, 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.

The selection of panel firms required additional adjustments to account for varying probabilities of selection between fresh and panel sample universes. For additional information on this methodology, please refer to Enterprise Survey documentation of weighting methodology.

34. Appendix C shows the cell weights for registered establishments in El Salvador.

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

35. 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.

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

---

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

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

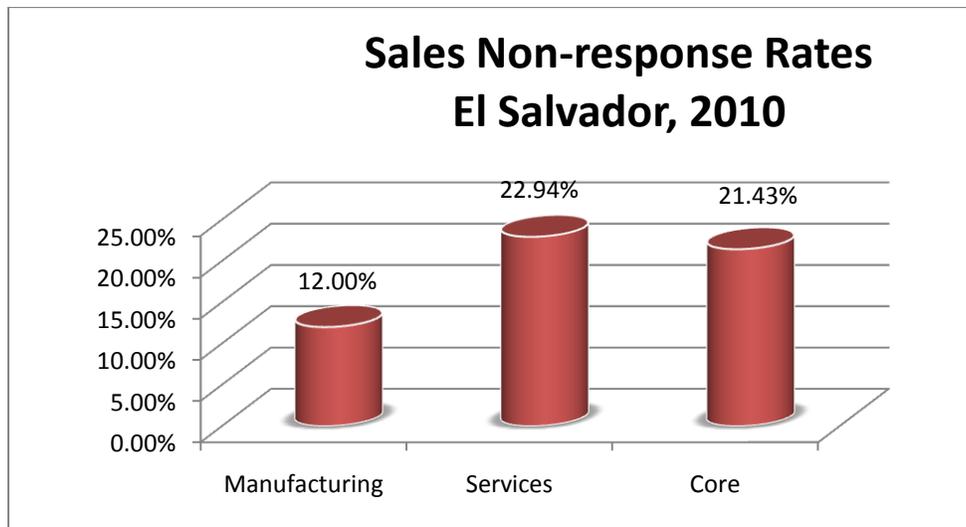
<sup>7</sup> Note that weighted OLS in Stata using the command `regress` with the option of weights will estimate wrong standard errors. Using the Stata survey specific commands `svy` will provide appropriate standard errors.

37. From a more general approach, if the regressions are descriptive of the population then weights should be used. The estimated model can be thought of as the relationship that would be expected if the whole population were observed.<sup>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

38. 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.

39. 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).



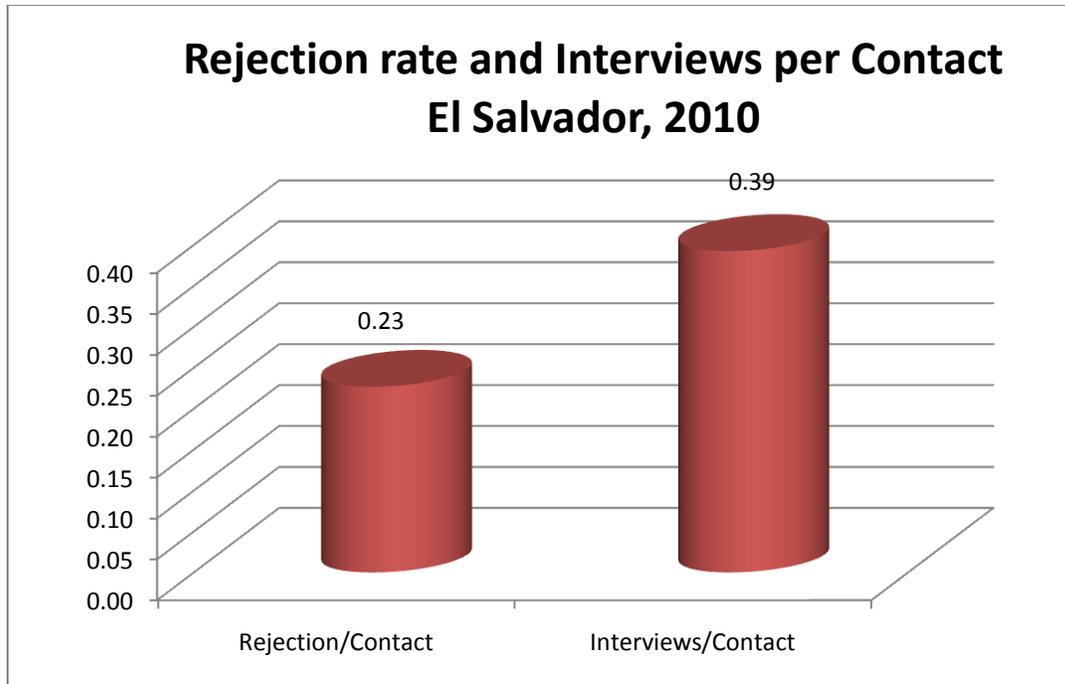
40. Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics)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

---

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

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.

41. As the following graph shows, the number of realized interviews per contacted establishment was 0.39<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 0.23.



42. 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 El Salvador. 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.

---

<sup>9</sup> The estimate is based on the total no. of firms contacted including ineligible establishments.

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 Fresh:

	<b>ELIGIBLES</b>	
Eligible	1. Eligible establishment (Correct name and address)	170
Eligible	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	0
Eligible	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	0
Eligible	4. Eligible establishment (Moved and traced)	3
		0
Ineligible	5. The establishment has less than 5 permanent full time employees	1
Ineligible	6 The firm discontinued businesses	8
Ineligible	7. Not a business: Private household	2
Ineligible	8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	1
Ineligible	151 Out of target - outside the covered regions	1
Ineligible	152. Out of target - moved abroad	0
Unobtainable	91. No reply after having called in different days of the week and in different business hours	48
Unobtainable	92. Line out of order	10
Unobtainable	93. No tone	0
Unobtainable	10. Answering machine	1
Unobtainable	11. Fax line- data line	2
Unobtainable	12. Wrong address/ moved away and could not get the new references	58
	13. Refuses to answer the screener	123
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	47
	<b>Total</b>	<b>475</b>

### Response Outcomes Fresh:

<b>Target</b>	
Complete interviews	168
Incomplete interviews	1
Elegible in process	4
Refusals	0
Out of target	13
Impossible to contact	119
Refusal to the Screener	123

428

### Status Codes Panel:

ELIGIBLES		
Eligible	1. Eligible establishment (Correct name and address)	226
Eligible	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	1
Eligible	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	1
Eligible	4. Eligible establishment (Moved and traced)	1
Eligible	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	0
Ineligible	616 The firm discontinued businesses - (Establishment went bankrupt)	6
Ineligible	618 The firm discontinued businesses - (Original establishment disappeared and is now a different firm)	2
Ineligible	619 The firm discontinued businesses - (Establishment was bought out by another firm)	2
Ineligible	620 The firm discontinued businesses - (It was impossible to determine for what reason)	2
Ineligible	621 The firm discontinued businesses - (Other: SPECIFY in COMMENTS)	1
Ineligible	7. Not a business: Private household	0
Ineligible	8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	0
Ineligible	151 Out of target - outside the covered regions	0
Ineligible	152. Out of target - moved abroad	0
Unobtainable	91. No reply after having called in different days of the week and in different business hours	65
Unobtainable	92. Line out of order	20
Unobtainable	93. No tone	5
Unobtainable	10. Answering machine	2
Unobtainable	11. Fax line- data line	1
Unobtainable	12. Wrong address/ moved away and could not get the new references	24
	13. Refuses to answer the screener	83
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	0
	Total	442

### Response Outcomes Panel:

Panel	
Complete interviews	192
Incomplete interviews	2
Elegible in process	33
Refusals	2
Out of target	13
Impossible to contact	117
Refusal to the Screener	83

442

## Appendix B

### Universe Estimate, El Salvador :

#### Universe

Region name	Employees	Manufacturing	Retail	Other Services	Grand Total
Rest of Country	5-19	255	847	1062	2164
	20-99	177	152	353	682
	100+	133	23	96	252
Rest of Country Total		565	1022	1511	3098
San Salvador	5-19	403	722	1517	2642
	20-99	267	212	582	1061
	100+	135	32	135	302
San Salvador Total		805	966	2234	4005
Grand Total		1370	1988	3745	7103

## Appendix C

### Strict Cell Weights El Salvador :+

#### Panel

Average Strict Cell Weights, Panel firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	2.80	1.39	1.53
	Medium	3.50	1.15	1.34
	Large	2.12	1.24	1.73
Regional Total				
Rest of the country	Small	3.12	1.36	1.52
	Medium	3.00	1.51	1.43
	Large	1.91	1.30	1.68

#### Fresh

Strict Cell Weights, Fresh Firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	16.10	35.68	61.95
	Medium	13.87	9.42	22.89
	Large	9.10	1.04	4.52
Regional Total				
Rest of the country	Small	8.97	25.80	47.74
	Medium	9.25	10.70	24.46
	Large	6.91	1.02	3.24

## Weak Cell Weights El Salvador :

### Panel

Average Weak Cell Weights, Panel firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	6.00	2.85	2.22
	Medium	6.93	2.48	1.94
	Large	2.71	2.20	2.82
Regional Total				
Rest of the country	Small	7.50	1.85	4.33
	Medium	4.09	1.83	2.50
	Large	5.57	2.11	2.05

### Fresh

Weak Cell Weights, Fresh Firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	43.99	100.99	147.57
	Medium	33.38	23.56	48.33
	Large	18.79	2.22	8.14
Regional Total				
Rest of the country	Small	27.27	81.24	126.50
	Medium	24.93	29.87	57.43
	Large	15.87	2.43	6.48

## Median Cell Weights El Salvador :

### Panel

Average Median Cell Weights, Panel firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	3.74	2.16	1.77
	Medium	4.36	1.85	1.49
	Large	2.29	2.00	2.28
Regional Total				
Rest of the country	Small	4.12	1.36	1.52
	Medium	3.09	1.69	1.73
	Large	3.43	1.78	1.81

### Fresh

Median Cell Weights, Fresh Firms

Location	Firm Size	Manuf.	Retail	Services
San Salvador	Small	25.86	65.97	97.54
	Medium	23.68	18.51	38.31
	Large	14.39	1.89	7.01
Rest of the country	Small	15.10	49.98	78.75
	Medium	16.55	22.04	42.88
	Large	11.44	1.95	5.25

## Appendix D

### Strict Universe Estimates

Strict Universe Estimates (Fresh + Panel)

Location	Firm Size	<b>Manuf.</b>	<b>Retail</b>	<b>Services</b>	<b>Grand Total</b>
San Salvador	Small	268.98	255.17	677.85	1202.00
	Medium	132.22	147.81	154.84	434.87
	Large	99.77	35.60	176.98	312.35
Regional Total		500.97	438.58	1009.67	1949.23
Rest of the country	Small	96.76	278.73	338.75	714.23
	Medium	93.14	70.08	132.28	295.50
	Large	87.18	19.23	40.93	147.33
Regional Total		277.08	368.04	511.95	1157.07
Grand Total		778.05	806.62	1521.63	3106.30

## Weak Universe Estimates

Weak Universe Estimates (Fresh + Panel)

Location	Firm Size	Manuf.	Retail	Services	Grand Total
San Salvador	Small	684.79	718.54	1574.40	2977.74
	Medium	297.25	345.87	315.16	958.27
	Large	177.52	78.08	360.28	615.89
Regional Total		1159.57	1142.49	2249.84	4551.90
Rest of the country	Small	278.16	846.93	898.53	2023.62
	Medium	203.02	169.44	304.64	677.11
	Large	215.79	38.35	72.63	326.77
Regional Total		696.98	1054.71	1275.81	3027.50
Grand Total		1856.55	2197.20	3525.65	7579.40

## Median Universe Estimates

Median Universe Estimates (Fresh + Panel)

Location	Firm Size	<b>Manuf.</b>	<b>Retail</b>	<b>Services</b>	<b>Grand Total</b>
San Salvador	Small	416.39	469.61	1069.82	1955.82
	Medium	203.08	270.95	249.26	723.29
	Large	139.71	62.13	274.83	476.66
Regional Total		759.18	802.68	1593.91	3155.77
Rest of the country	Small	153.78	526.68	555.81	1236.28
	Medium	140.45	128.82	226.51	495.77
	Large	148.65	31.47	59.92	240.04
Regional Total		442.88	686.97	842.24	1972.09
Grand Total		1202.06	1489.65	2436.15	5127.86

**Appendix E**

**Original Sample Design, El Salvador :**

Region name	Employees	Manufacturing	Retail	Other Services	Grand Total
Rest of Country	5-19	16	22	16	54
	20-99	16	17	15	48
	100+	19	16	16	51
Rest of Country Total		51	55	47	153
San Salvador	5-19	24	18	24	66
	20-99	24	23	25	72
	100+	21	24	24	69
San Salvador Total		69	65	73	207
Grand Total		120	120	120	360

**Completed Interviews, El Salvador :**

Region name	Employees	Manufacturing	Retail	Other Services	Grand Total
Rest of Country	5-19	14	25	17	56
	20-99	20	11	17	48
	100+	21	12	13	46
Rest of Country Total		55	48	47	150
San Salvador	5-19	26	16	34	76
	20-99	20	26	23	69
	100+	24	19	22	65
San Salvador Total		70	61	79	210
Grand Total		125	109	126	360

## Appendix F

### Local Agency team involved in the study:

*NOTE: Fieldwork in El Salvador was completed by two local agencies*

### Local Agency 1 team involved in the study:

Local Agency	Mercaplan
Enumerators involved:	13
Other staff involved:	1- Quality Control Validation

### Local Agency 2 team involved in the study:

Local Agency	Data Research, S.A. de C.V.
Enumerators involved:	10
Other staff involved:	1

### Sample Frame:

Characteristic of sample frame used:	Business Directory (Directorio de Empresas de bases de solvencias)
Source:	General Directorate of Statistics and Censuses (Dirección General de Estadísticos y Censos). 2008
Comments on the quality of sample frame:	There were various problems with the sample frame. We found several companies with less than 5 employees. Others were repeated or with the wrong phone number or address.
Year and organism who conducted the last economic census	General Directorate of Statistics and Censuses (Dirección General de Estadística y Censos 2008)
Other sources for companies statistics	

### Sectors included in the Sample:

Original Sectors	Manufacturing, Retail and Core (services)
Added Sectors	None

**Sample:**

Comments/ problems on sectors and regions selected in the sample. Agency 1	The panel sample was based on a study conducted 4 years ago, due to that, we found that many establishments were impossible to contact. In many cases the addresses and phone numbers were not correct.  We also consider that the substitutes weren't enough in order to achieve an interview.
Agency 2	There were no problems with sectors or regions selected.
Comments on the response rate. Agency 1	None
Agency 2	The response rate obtained could improve if social and economic aspects of the country will improve in the future.
Comments on the sample design. Agency 1	Probably it would have been more effective not to work with a completely random sample, especially for those establishments which were contacted in the previous enterprise survey.
Agency 2	The sample design was highly appropriate for our country's characteristics
Other comments. Agency 1	None
Agency 2	None

**Fieldwork:**

Date of Fieldwork. Agency 1	July 8 <sup>th</sup> – December 6 <sup>th</sup> 2010
Agency 2	December 15 <sup>th</sup> 2010 - April 19 <sup>th</sup> 2011
Problems found during fieldwork. Agency 1	The main issue was distrust. Generally, respondents were reluctant to give information (especially financial), mainly because of insecurity issues and fear of future extortions. Another problem was the length of the questionnaire. The respondents, owners/managers, usually did not have time to answer the interview.
Agency 2	Most companies have the policy not to provide financial information or any information they consider could expose them to the criminal and violent environment that prevails in the country
Other observations.	None

Agency 1	
Agency 2	December is not a good month for conducting fieldwork since is the year's main commercial time for many companies, and other companies use that time for personnel vacations.

**Questionnaires:**

Problems for the understanding of questions (write question number). Agency 1	None
Agency 2	There were no questions that were difficult to understand.
Problems found in the navigability of questionnaires. Agency 1	None
Agency 2	There were no problems found in the navigability of questionnaires.
Comments on questionnaires length. Agency 1	In many occasions interviewers complained about the length of the questionnaire. Usually the respondents were very busy; this affected the survey's quality since they tended to be distracted by phone calls and other matters.
Agency 2	Questionnaire's length always presented problems due to the interviewee's position.
Suggestions or other comments on the questionnaire. Agency 1	None
Agency 2	As stated above, questionnaires should be shorter.

**Database**

Comments on the data map. Agency 1	None
Agency 2	The data map was qualified as excellent.
Comments on the data processing. Agency 1	None
Agency 2	Data processing was highly facilitated given its structure

## Country situation

General aspects of economic, political or social situation of the country that could affect the results of the survey. Agency 1	There were multiple events which could have affected the survey results, especially fiscal reform and municipal tax increase.
Agency 2	In recent years, El Salvador has been going through a social situation that has created a high rate of internal violence. This violence makes everyone very reluctant to provide any type of economic information.
Relevant country events occurred during fieldwork. Agency 1	During fieldwork, we went through a bus strike.
Agency 2	Fieldwork was started at the beginning of the Christmas season, event that highly interfered with its development for well known reasons.
Other aspects. Agency 1	None
Agency 2	None