

The Costa Rica 2010 Enterprise Surveys Data Set

I. Introduction

1. This document provides additional information on the data collected in Costa Rica between July 2010 and June 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 Costa Rica was selected using stratified random sampling, following the methodology explained in the *Sampling Note*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/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.)

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 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 3 manufacturing industries (isic codes 15, 26-29, and Other Manufacturing), 2 service industry (retail and hotels/restaurants), and 1 residual services sector as defined in the sampling manual. Each of the manufacturing industries and the retail sector had a target each of 120 interviews. Hotels/restaurants and other services each had a target of 60 interviews each.

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 Jose and Central Valley 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 Costa Rica the local subcontractor was Borge y Asociados.

9. For Costa Rica, two sample frames were used. The first was supplied by the World Bank and consists of enterprises interviewed in Costa Rica in 2006. The World Bank required that attempts should be made to re-interview establishments responding to the Costa Rica 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 Census of businesses and commercial establishments from 2010 created by the National Institute of Statistics (INEC).

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

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	9	10	12		11		42
	20-99	6	6	11		6		29
	100+			3		1		4
Rest of Country Total		15	16	26		18		75
San Jose and Central Valley	5-19	6	23	64		10		103
	20-99	9	16	39		6		70
	100+	8	7	20		1		36
San Jose and Central Valley Total		23	46	123		17		209
Grand Total		38	62	149		35		284

Sample Frames

Source: Directorio de establecimientos del INEC (Instituto Nacional de Estadísticas y Censos), 2010.

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	60	30	68	478	628	553	1817
	20-99	23	5	24	87	152	141	432
	100+	21	3	6	11	22	16	79
Rest of Country Total		104	38	98	576	802	710	2328
San Jose and Central Valley	5-19	211	134	328	1178	657	1784	4292
	20-99	106	89	289	270	145	944	1843
	100+	65	28	159	67	52	301	672
San Jose and Central Valley Total		382	251	776	1515	854	3029	6807
Grand Total		486	289	874	2091	1656	3739	9135

10. The two sample frames were then used for the selection of a sample with the aim of obtaining interviews with 600 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 5.69% (88 out of 1546 establishments)⁴.

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*.

⁴ 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 (11, 16 and 18) 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. Variables a17x gives interviewer comments, including problems that occurred during an interview and extraordinary circumstances which could influence results. Please note that sometimes this variable is removed due to privacy issues.

V. Universe Estimates

22. Universe estimates for the number of establishments in each cell in Costa Rica 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 Costa Rica 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*.

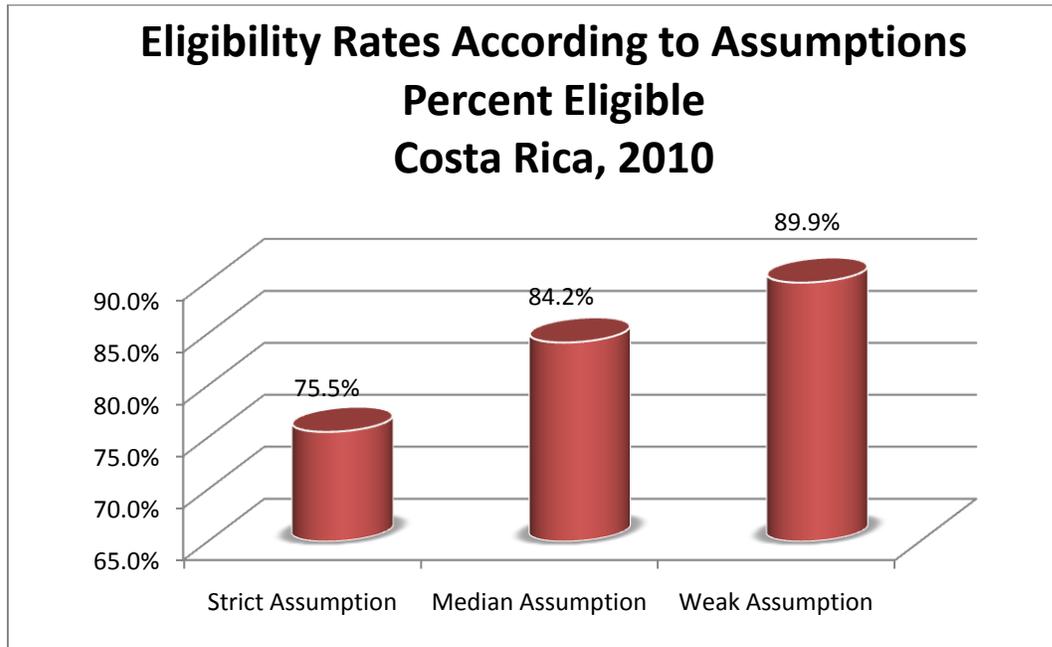
Median eligibility = (Sum of the firms with codes 1,2,3,4,16,10,11, & 13) / Total

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 Costa Rica 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.)⁵

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 line⁶, 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 Costa Rica.

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

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

⁶ For the surveys that implemented a screener over the phone.

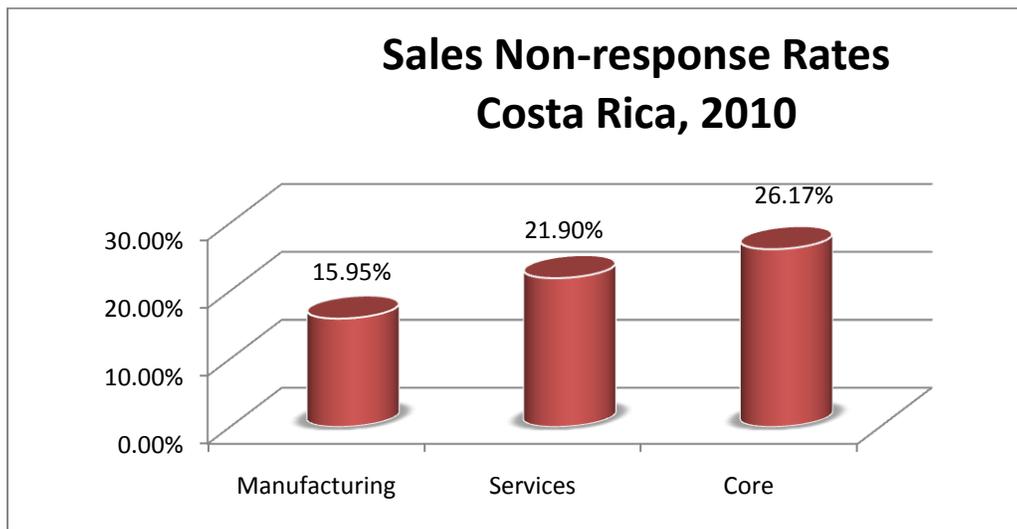
estimates but also design-unbiased estimates (see also Cochran, 1977, pp 200 who favors the used of weighted OLS for a common population coefficient.)⁷

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.⁸ 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).

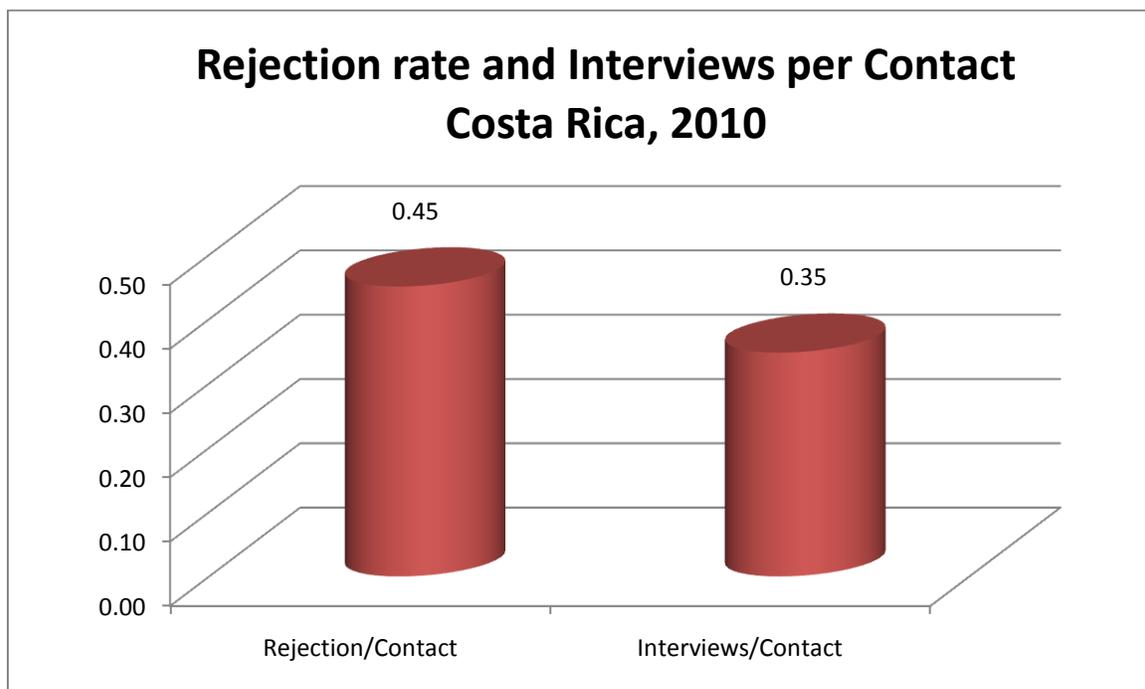


⁷ Note that weighted OLS in Stata using the command regress with the option of weights will estimate wrong standard errors. Using the Stata survey specific commands svy will provide appropriate standard errors.

⁸ 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.

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 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 0.35⁹. 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 0.45.



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 Costa Rica. All enterprise surveys suffer from these shortcomings, but in very few cases they have been made explicit.

References:

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

Cochran, William G., Sampling Techniques, 1977.

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

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

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

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

Appendix A

Status Codes Fresh:

	ELIGIBLES	
Eligible	1. Eligible establishment (Correct name and address)	977
Eligible	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	3
Eligible	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	1
Eligible	4. Eligible establishment (Moved and traced)	0
		0
Ineligible	5. The establishment has less than 5 permanent full time employees	24
Ineligible	6 The firm discontinued businesses	6
Ineligible	7. Not a business: Private household	9
Ineligible	8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	39
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	44
Unobtainable	92. Line out of order	12
Unobtainable	93. No tone	0
Unobtainable	10. Answering machine	0
Unobtainable	11. Fax line- data line	1
Unobtainable	12. Wrong address/ moved away and could not get the new references	7
	13. Refuses to answer the screener	96
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	68
	Total	1287

Response Outcomes Fresh:

Target	
Complete interviews	429
Incomplete interviews	0
Elegible in process	70
Refusals	490
Out of target	78
Impossible to contact	64
Refusal to the Screener	96

1227

Status Codes Panel:

	ELIGIBLES	
Eligible	1. Eligible establishment (Correct name and address)	183
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)	3
Eligible	4. Eligible establishment (Moved and traced)	0
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)	1
Ineligible	618 The firm discontinued businesses - (Original establishment disappeared and is now a different firm)	1
Ineligible	619 The firm discontinued businesses - (Establishment was bought out by another firm)	0
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	3
Ineligible	8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	2
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	16
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	0
	13. Refuses to answer the screener	33
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	0
	Total	259

Response Outcomes Panel:

Panel	
Complete interviews	109
Incomplete interviews	0
Elegible in process	0
Refusals	78
Out of target	10
Impossible to contact	29
Refusal to the Screener	33

259

Appendix B

Universe Estimate, Costa Rica:

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	51	20	56	478	617	553	1775
	20-99	17	3	13	87	146	141	407
	100+	21	3	6	11	21	16	78
Rest of Country Total		89	26	75	576	784	710	2260
San Jose and Central Valley	5-19	205	111	264	1178	647	1784	4189
	20-99	97	83	250	270	139	944	1783
	100+	59	24	139	67	51	301	641
San Jose and Central Valley Total		361	218	653	1515	837	3029	6613
Grand Total		450	244	728	2091	1621	3739	8873

Appendix C

Strict Cell Weights Costa Rica:

Panel Collapsed Cell Weights

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	1.01	1.22	6.59		1.67	
	20-99	1.97	1.31	1.78		3.78	
	100+			2.20		1.00	
San Jose and Central Valley	5-19	1.40	1.49	2.04		1.67	
	20-99	1.48	2.13	1.91		1.00	
	100+	3.75	1.44	2.48			

Fresh Strict Cell Weights, Fresh Firms, Collapsed Cell Weights

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	10.32	17.85	21.81	38.10	74.42	84.98
	20-99	4.22		2.98	8.59	24.33	
	100+	3.08	2.82	1.84	5.14	5.25	5.13
San Jose and Central Valley	5-19	4.21	3.51	13.79	32.01	74.42	99.91
	20-99	3.05	2.32	8.98	7.26	25.50	61.27
	100+	2.67	4.41	3.42	2.59	3.16	14.74

Weak Cell Weights Costa Rica:

Panel

Collapsed cell Weights

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	1.57	2.36	11.37		2.63	
	20-99	2.56	2.02	2.55		5.71	
	100+			2.85		1.00	
San Jose and Central Valley	5-19	1.82	2.35	2.95		2.63	
	20-99	1.60	2.74	2.32		1.20	
	100+	3.66	1.73	2.70			

Fresh

Weak Cell Weights, Fresh Firms

Collapsed Cell Weights

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	12.13	21.11	25.76	45.39	89.69	112.30
	20-99	4.25		2.95	8.70	24.33	
	100+	3.15	2.80	1.88	5.34	5.11	5.33
San Jose and Central Valley	5-19	5.50	4.48	17.96	42.42	89.69	132.34
	20-99	3.46	2.57	10.12	8.34	27.08	68.72
	100+	3.04	5.00	3.88	2.99	3.37	17.27

Median Cell Weights Costa Rica:

Panel

Collapsed Cell Weights

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	1.24	1.77	8.03		2.10	
	20-99	2.33	1.72	2.08		4.72	
	100+			2.30		1.00	
San Jose and Central Valley	5-19	1.65	2.02	2.41		2.10	
	20-99	1.68	2.69	2.21		1.14	
	100+	3.78	1.71	2.52			

Fresh

Median Cell Weights, Fresh Firms

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services
Rest of Country	5-19	10.88	18.67	23.05	40.09	79.88	97.32
	20-99	4.22		2.88	8.59	24.33	
	100+	3.15	2.78	1.89	5.24	5.10	5.26
San Jose and Central Valley	5-19	4.89	3.98	16.12	37.07	79.88	114.61
	20-99	3.36	2.49	9.94	7.98	26.37	66.72
	100+	3.01	4.90	3.87	2.91	3.33	16.63

Appendix D

Strict Universe Estimates

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	36	58	94	305	227	510	1229
	20-99	21	5	8	86	150	0	270
	100+	22	3	6	10	22	15	78
Rest of Country Total		78	66	108	401	398	525	1577
San Jose and Central Valley	5-19	160	115	169	928	754	1099	3225
	20-99	87	78	230	232	132	980	1740
	100+	58	53	105	57	44	236	553
San Jose and Central Valley Total		305	246	504	1217	931	2315	5519
Grand Total		383	312	612	1618	1329	2840	7096

Weak Universe Estimates

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	44	73	114	363	274	674	1543
	20-99	22	8	11	87	152	0	279
	100+	22	3	7	11	21	16	80
Rest of Country Total		88	84	132	461	447	690	1902
San Jose and Central Valley	5-19	209	151	227	1230	913	1456	4185
	20-99	98	88	262	267	141	1100	1956
	100+	65	60	119	66	47	276	633
San Jose and Central Valley Total		372	300	607	1563	1101	2832	6775
Grand Total		461	383	739	2023	1549	3522	8676

Median Universe Estimates

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	39	63	100	321	244	584	1351
	20-99	22	7	9	86	151	0	274
	100+	22	3	6	10	21	16	79
Rest of Country Total		82	73	115	417	416	600	1703
San Jose and Central Valley	5-19	186	134	198	1075	811	1261	3665
	20-99	96	86	256	256	138	1068	1898
	100+	65	59	117	64	47	266	618
San Jose and Central Valley Total		347	278	571	1395	996	2594	6181
Grand Total		429	351	687	1812	1412	3194	7884

Appendix E

Original Sample Design, Costa Rica:

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	9	9	6	11	8	4	47
	20-99	7	3	3	10	9	2	34
	100+	11	3	6	7	9	9	45
Rest of Country Total		27	15	15	28	26	15	126
San Jose and Central Valley	5-19	30	42	31	28	9	13	153
	20-99	31	49	33	30	8	15	166
	100+	32	14	41	34	17	17	155
San Jose and Central Valley Total		93	105	105	92	34	45	474
Grand Total		120	120	120	120	60	60	600

Completed Interviews, Costa Rica:

Region name	Employees	15	26,27,28&29	Other Manufacturing	52	55	Other Services	Grand Total
Rest of Country	5-19	9	8	6	9	5	4	41
	20-99	6	3	5	10	9	3	36
	100+	6	1	1	1	3	2	14
Rest of Country Total		21	12	12	20	17	9	91
San Jose and Central Valley	5-19	34	41	35	29	16	8	163
	20-99	34	29	44	32	15	19	173
	100+	24	16	24	22	9	16	111
San Jose and Central Valley Total		92	86	103	83	40	43	447
Grand Total		113	98	115	103	57	52	538

Appendix F

Local Agency team involved in the study:

Local Agency	Borge y Asociados
Enumerators involved:	21
Other staff involved:	4

Sample Frame:

Characteristic of sample frame used:	Census of businesses and commercial establishments
Source:	INEC
Year:	2004 and subsequent updates (last update – October 2010)
Comments on the quality of sample frame:	None
Year and organism who conducted the last economic census	INEC
Other sources for companies statistics	Business Phone Directory

Sectors included in the Sample:

Original Sectors	15; 26,27,28&29; Other Manufacturing; Retail; Hotels and Restaurants; Other Services
Added Sectors	None

Sample:

Comments/ problems on sectors and regions selected in the sample:	At least 10% of the Panel Sample did not match the sample surveyed in 2005 .
Comments on the response rate:	The hard work dedicated to contacting establishments made response rate higher.
Comments on the sample design:	The process of drawing new sample when the issued contacts were exhausted produced significant delay.

Fieldwork:

Date of Fieldwork	July 2010 – June 2011
Problems found during fieldwork:	Natural disasters, holidays and changes in the geographical location.
Other observations:	None

Questionnaires:

Problems for the understanding of questions (write question number)	Question D.30: ¿Cuál es el grado de obstáculo que representa el transporte para el funcionamiento actual de este establecimiento? In our view, the sentence above should have been rephrased as “ el acceso al transporte ”.
Problems found in the navigability of questionnaires.	Difficulties in questions related with the public services failures (water and electricity)
Comments on questionnaires length:	Managers thought the questionnaire was too long
Suggestions or other comments on the questionnaire:	Most of the interviewees suggested to include more questions regarding the amount of time to obtain credits.

Database

Comments on the data map	None
Comments on the data processing	None

Country situation

General aspects of economic, political or social situation of the country that could affect the results of the survey:	None
Relevant country events occurred during fieldwork:	None
Other aspects:	None