The Madagascar 2013 Enterprise Surveys Data Set

I. Introduction

1.   This document provides additional information on the data collected in Madagascar between November 2013 and May 2014 under, an initiative of the World Bank. ***This document provides information only on the portion of the survey that is comparable to other Enterprise Surveys. Portions of this survey dealing with Micro-firms (less than 5 employees) and industries outside of the Enterprise Survey standard universe of interest (e.g. Agriculture & Extractive Industries) are referenced by this document as they were part of an integrated survey strategy. Observations from these portions of the survey work are excluded from indicator calculations shown on www.enterprisesurveys.org and from the analysis that follow in this report.***

As part of its strategic goal of building a climate for investment, job creation, and sustainable growth, the World Bank has promoted improving business environments as a key strategy for development, which has led to a systematic effort in collecting enterprise data across countries. The Enterprise Surveys (ES) are an ongoing World Bank project in collecting both objective data based on firms' experiences and enterprises' perception of the environment in which they operate.

The Enterprise Surveys currently cover over 130,000 firms in 135 countries, of which 121 have been surveyed following a standard methodology. This allows for better comparisons across countries and across time. Data are used to create statistically significant business environment indicators that are comparable across countries. The Enterprise Surveys are also used to build a panel of enterprise data that will make it possible to track changes in the business environment over time and allow, for example, impact assessments of reforms.

The report outlines and describes the sampling design of the data, the data set structure as well as additional information that may be useful when using the data, such as information on non-response cases and the appropriate use of the weights.

II. Sampling Structure

2.   The sample for Madagascar was selected using stratified random sampling, following the methodology explained in the *Sampling Manual*[1]. Stratified random sampling[2] was preferred over simple random sampling for several reasons[3]:

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. *Additional sectors in extractive industries and agriculture were included as part of the Madagascar sample but are not used to produce indicators for the www.enterprisesurveys.org website and are not covered by this implementation report.*

---

[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

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.

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 four manufacturing industries (food, textiles and garments, chemicals and plastics, other manufacturing) and two service sectors (retail and other services).

5.   For the Madagascar Enterprise Survey (ES), 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). The micro sample consists of firms with 1 to 4 employees. A significant portion of the sample frame did not contain size information and so a 5th size category 'undefined' was also used. *Firms included as part of the Madagascar sample but found to have fewer than 5 full-time-equivalent employees are not used to produce indicators for the www.enterprisesurveys.org website and are not covered by this implementation report.*

6.   Regional stratification for the Madagascar ES was defined in eight regions:
- Analamanga
- Anôsy
- Atsimo Andrefana
- Atsinanana
- Boeny
- Diana
- Nosy be
- Vakinankaratra

III. Sampling implementation

7.   Given the stratified design, sample frames containing a complete and updated list of establishments as well as information on all stratification variables (number of employees, industry, and region) are required to draw the sample.

8.   The Malagasy firm HERMES was hired to implement the Madagascar survey

9.   For the Madagascar ES, a single sample frame used from Institut National de la Statistique de Madagascar (l'INSTAT).

Madagascar, Sample Frame

REGION	Size	Raw material processing manufacturing	Food & Beverages	Textiles & Garments	Other Manufacturing	Hospitality &'Tourism	Other services	TOTAL
ANALAMANGA	unknown	48	38	41	155	226	1601	2109
ANALAMANGA	Small	42	50	62	207	299	1734	2394
ANALAMANGA	Medium	14	11	37	51	69	237	419
ANALAMANGA	Large	6	11	58	15	12	92	194
ANOSY	unknown	0	2	0	2	8	33	45
ANOSY	Small	0	3	0	2	12	23	40
ANOSY	Medium	0	0	0	0	1	2	3
ANOSY	Large	0	0	0	0	0	2	2
ANTSINANANA	unknown	0	0	0	1	1	5	7
ANTSINANANA	Small	6	15	2	25	143	452	643
ANTSINANANA	Medium	2	3	0	6	14	50	75
ANTSINANANA	Large	1	1	0	1	6	6	15
ATSIMO ANDREFANA	unknown	0	1	0	2	44	73	120
ATSIMO ANDREFANA	Small	0	4	0	4	43	62	113
ATSIMO ANDREFANA	Medium	0	0	0	1	7	4	12
ATSIMO ANDREFANA	Large	0	0	0	0	0	3	3
BOENI	unknown	0	0	0	0	0	9	9
BOENI	Small	5	14	1	29	76	294	419
BOENI	Medium	2	5	2	6	5	29	49
BOENI	Large	1	1	1	0	3	2	8
DIANA	unknown	1	4	0	4	49	117	175
DIANA	Small	3	4	0	6	33	71	117
DIANA	Medium	0	1	0	1	8	34	44
DIANA	Large	0	1	0	0	2	2	5
Nosy Be	unknown	1	0	0	3	33	30	67
Nosy Be	Small	1	0	0	8	38	28	75
Nosy Be	Medium	0	0	0	0	14	6	20
Nosy Be	Large	0	0	0	0	5	0	5
VAKINAKARATRA	unknown	0	0	0	0	0	1	1
VAKINAKARATRA	Small	9	39	8	60	67	431	614
VAKINAKARATRA	Medium	1	4	1	2	8	28	44
VAKINAKARATRA	Large	0	5	3	1	0	3	12
TOTAL		143	217	216	592	1226	5464	7858

10. The sample design for the Madagascar Enterprise Survey was generated with the aim of obtaining interviews at 589 establishments. Establishments with undefined size were included as part of this single sample frame for Madagascar in order to ensure a representative sample. Size information collected during the survey process was used to categorize these firms into their respective bins.

11.   The quality of the frame was assessed at the onset of the project through visits to a random subset of firms and local contractor knowledge. The sample frame was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc. The local contractor had to screen the contacts by visiting them which resulted in slow fieldwork in many cases.

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 37% (1182 out of 3183 establishments)[4]. Breaking down by stratified industries, the following sample targets were achieved (using a4a and a6a):

---

[4] Based on out of target contacts and impossible to contact establishments

Achieved sample

REGION	Size	Raw material processing manufacturing	Food & Beverages	Textiles & Garments	Other Manufacturing	Hospitality &'Tourism	Other services	TOTAL
ANALAMANGA	Small	8	17	13	8	8	21	75
ANALAMANGA	Medium	7	5	13	13	6	17	61
ANALAMANGA	Large	11	8	29	9	3	6	66
ANOSY	Small	1	1		6	12	11	31
ANOSY	Medium		1			2	4	7
ANOSY	Large					1		1
ANTSINANANA	Small		1		4	1	8	14
ANTSINANANA	Medium	3	1		1	3	5	13
ANTSINANANA	Large	1				1	1	3
ATSIMO ANDREFANA	Small	3	14	1	14	23	10	65
ATSIMO ANDREFANA	Medium	1	1			7	2	11
ATSIMO ANDREFANA	Large		3					3
BOENI	Small	4	10	2	11		3	30
BOENI	Medium	1	2		1	1		5
BOENI	Large	2	2			1		5
DIANA	Small	1	2	1		31	4	39
DIANA	Medium		1			6	1	8
DIANA	Large					1		1
NOSY BE	Small	2		2	3	32	4	43
NOSY BE	Medium		1		3	9	4	17
NOSY BE	Large					1		1
VAKINAKARATRA	Small	1	6	1	9		7	24
VAKINAKARATRA	Medium	1			1	1	3	6
VAKINAKARATRA	Large			1	2			3
TOTAL		**47**	**74**	**65**	**85**	**150**	**111**	**532**

IV. Data Base Structure:

13.   The structure of the data base reflects the fact that 2 different versions of the survey instrument were used for all registered establishments. Questionnaires have common questions (*core* module) and respectfully additional manufacturing and retail specific questions. The eligible manufacturing industries have been surveyed using the **Manufacturing** questionnaire (includes the *core* module, plus manufacturing specific questions). Eligible services have been covered using the core module only (**Other Services** questionnaire). Each variation of the questionnaire is identified by the index variable, *a0*. *Note that extractive industries and agriculture observations are not used to compute indicators for the www.enterprisesurveys.org website and are not covered by this implementation report.*

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* (some exceptions apply due to comparability reasons). Variable names proceeded by a prefix *"MG" or "HO"* indicate questions specific to Madagascar, 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 region), *a6a* (sampling establishment's size), and *a4a* (sampling sector) contain the establishment's

classification into the strata chosen for each country using information from the sample frame. The strata were defined according to the guidelines described above.

16.   There are three levels of stratification: industry, size and region. Different combinations of these variables generate the strata cells for each industry/region/size combination. A distinction should be made between the variable a4a and d1a2 (industry expressed as ISIC rev. 3.1 code). The former gives the establishment's classification into one of the chosen industry-strata, whereas the latter gives the actual establishment's industry classification (four digit code) in the sample frame. *Note that it was ascertained after the completion of the project that the classification for a4a and a4b are based on company name and so there are a large number of firms for which these variables do not accurately reflect the product or service of the business. We recommend using variables d1a2 and d1a1x instead.*

17.   All of the following variables contain information from the sampling frame. They may not coincide with the reality of individual establishments as sample frames may contain inaccurate information. The variables containing the sample frame information are included in the data set for researchers who may want to further investigate statistical features of the survey and the effect of the survey design on their results.
   -*a2* is the variable describing sampling regions
   -*a6a*: coded using the same standard for micro, 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 variables for size (*l1*, *l6* and *l8*) that reflect more accurately the reality of each establishment. Advanced users are advised to use these variables for analytical purposes. 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.

20.   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.

21. Note that the fiscal years vary by firm as there is no standard for all firms in Madagascar. The start and end dates for the fiscal year for each firm can be found in the a20 variables in the dataset

V. Universe Estimates

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

22. Appendix B shows the overall estimates of the numbers of establishments in Madagascar based on the sample frame.

23. For some establishments where contact was not successfully completed during the screening process (because the firm has moved and it is not possible to locate the new location, for example), it is not possible to directly determine eligibility. Thus, different assumptions about the eligibility of establishments result in different adjustments to the universe cells and thus different sampling weights.

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

25. Strict assumption: eligible establishments are only those for which it was possible to directly determine eligibility. The resulting weights are included in the variable *wstrict*.

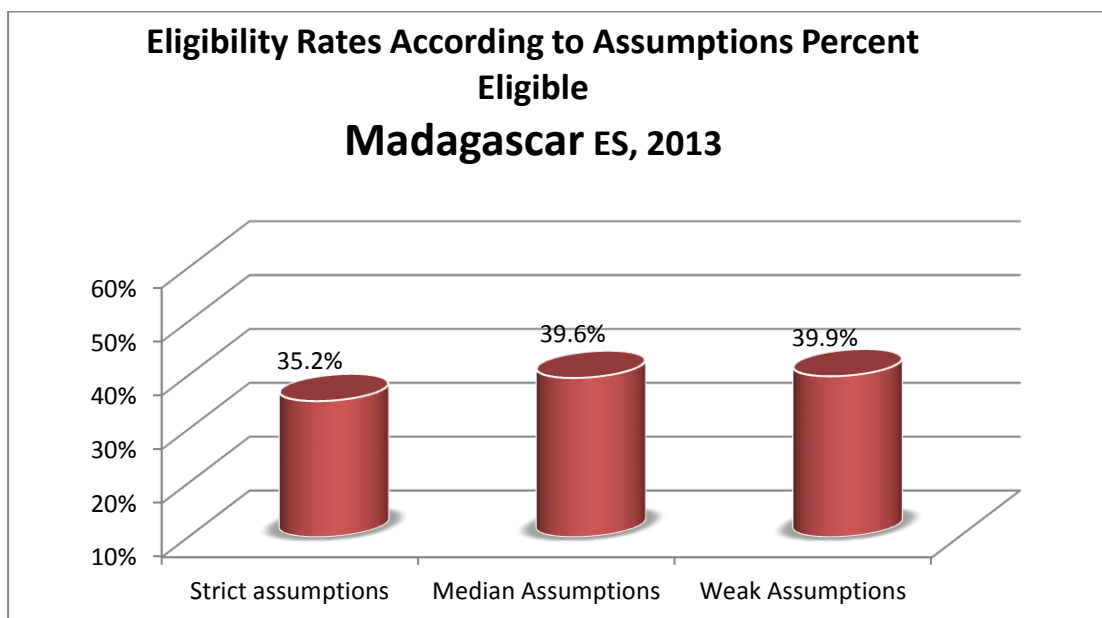*Strict eligibility = (Sum of the firms with codes 1,2,3,4,&16) / Total*

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

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

27. Weak assumption: in addition to the establishments included in points a and b, all establishments for which it was not possible to contact or that refused the screening questionnaire are assumed eligible. This definition includes as eligible establishments with dead or out of service phone lines, establishments that never answered the phone, and establishments with incorrect addresses for which it was impossible to find a new address. Under the weak assumption only observed non-eligible units are excluded from universe projections. The resulting weights are included in the variable *wweak*.

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

28. The indicators computed for the Enterprise Survey website use the median weights. The following graph shows the different eligibility rates calculated for firms in the sample frame under each set of assumptions.

Eligibility Rates According to Assumptions Percent Eligible
Madagascar ES, 2013

Assumption	Eligible
Strict assumptions	35.2%
Median Assumptions	39.6%
Weak Assumptions	39.9%

29. Universe estimates for the number of establishments in each industry-region-size cell in Madagascar 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.

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.

VI. 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 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[6], 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.

---

[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.

VII. Appropriate use of the weights

33.   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.

34.   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.)[7]

35.   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

36.   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.

37.   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 for both enterprise surveys (ES) reflect 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.

Sales Non-response Rates
Madagascar ES, 2013
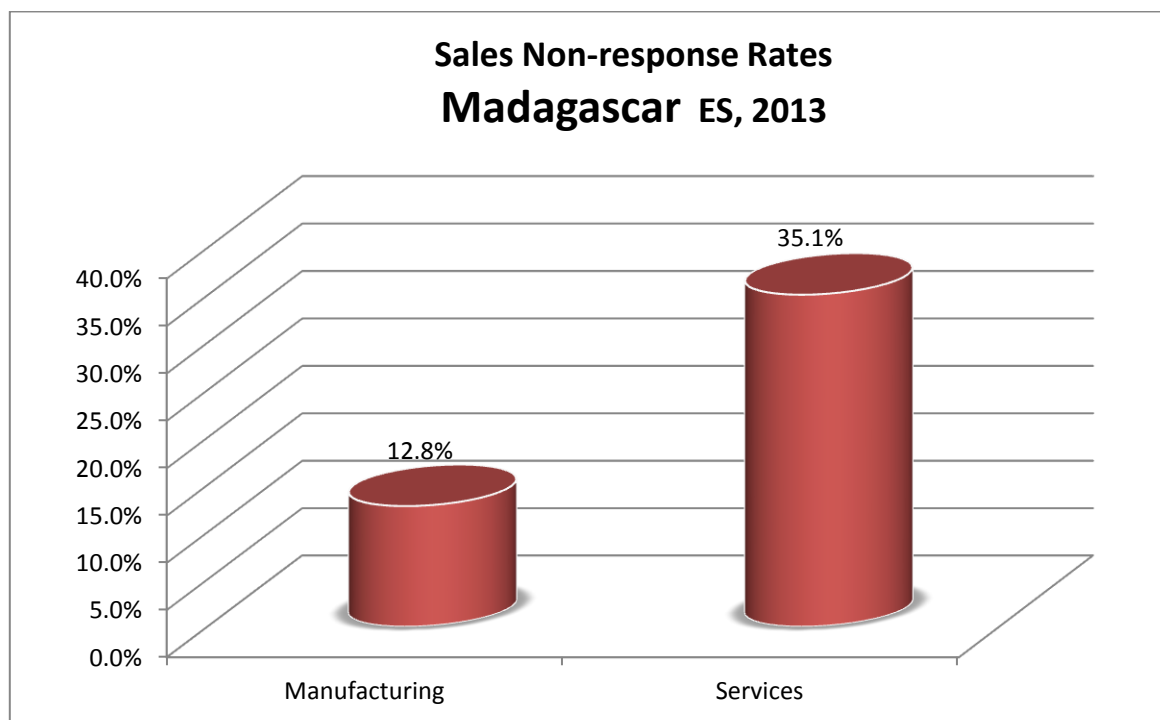
38.   Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) 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.

39.   As the following graph shows, the number of interviews per contacted establishments was 0.31[9]. 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 0.068.

Rejection rate and Interviews per Contact
Madagascar ES, 2013

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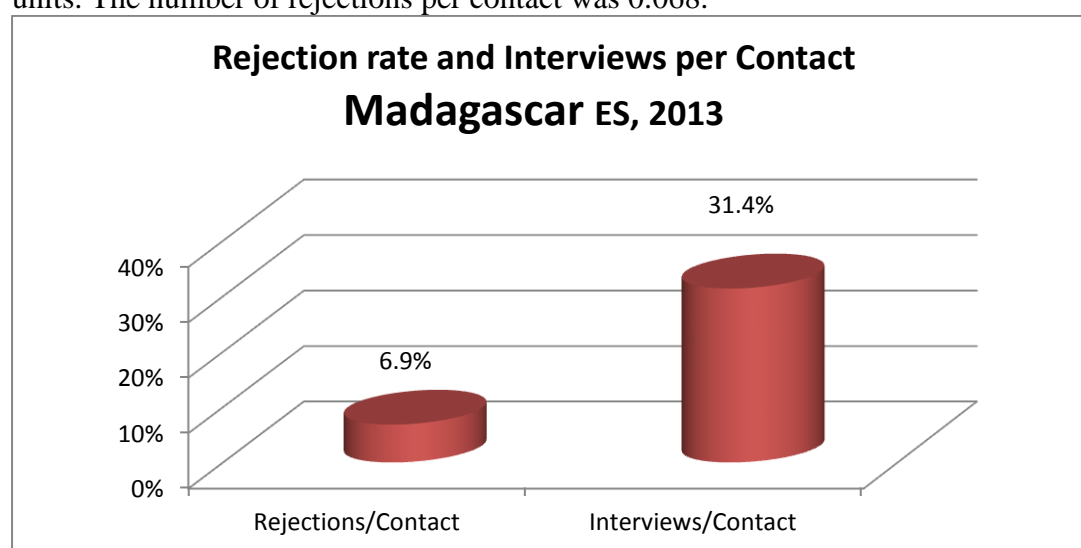
[9] The estimate is based on the total no. of firms contacted including ineligible establishments.

40.   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 Madagascar. 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. Samping: Design and Techniques, 1999.

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

Appendix A

Status Codes:

			Madagascar
		Sample Target	589
		Complete interviews (Total)	532
		Incomplete interviews	1
		Elegible in process	26
		Refusals	96
		Out of target	508
		Impossible to contact	739
		Ineligible - coop.	676
		Refusal to the Screener	136
		Total	2714

		Madagascar
Eligibles	1.Elegible establishment (Correct name and address)	1098
	2. Elegible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	5
	3. Elegible establishment (Different name but same address - the firm/establishment changed its name)	4
	4. Elegible establishment (Wrong address - the firm/establishmen has changed address and the address could be found)	11
Ineligibles	5. The establishment has less than 5 permanent full time employees	1
	6. The firm discontinued businesses	350
	7. Not a business: private household	146
	8. Ineligible activity: education, agriculture, finances, governments...	10
Unobtainable	91. No reply *(after having called in different days of the week and in different business hours)*	18
	92. Line out of order	6
	93. No tone	1
	94. Phone number does not exist	1
	10. Answering machine	0
	11. Fax line - data line	1
	12. Wrong address/ moved away and could not get the new references	712
	13. Refuses to answer the screener	136
	14. In process *(the establishment is being called/ is being contacted - previous to ask the screener)*	17
	151. Out of target - outside the covered regions, firm moved abroad	644
	152. Out of target - firm moved abroad	24
	153. Out of target - Not registered with SAT	8
	Total	3193

Appendix B   Universe Estimates, Madagascar:

REGION	Size	Raw material processing manufacturing	Food & Beverages	Textiles & Garments	Other Manufacturing	Hospitality &'Tourism	Other services	TOTAL
ANALAMANGA	unknown	48	38	41	155	226	1601	2109
ANALAMANGA	Small	42	50	62	207	299	1734	2394
ANALAMANGA	Medium	14	11	37	51	69	237	419
ANALAMANGA	Large	6	11	58	15	12	92	194
ANOSY	unknown	0	2	0	2	8	33	45
ANOSY	Small	0	3	0	2	12	23	40
ANOSY	Medium	0	0	0	0	1	2	3
ANOSY	Large	0	0	0	0	0	2	2
ANTSINANANA	unknown	0	0	0	1	1	5	7
ANTSINANANA	Small	6	15	2	25	143	452	643
ANTSINANANA	Medium	2	3	0	6	14	50	75
ANTSINANANA	Large	1	1	0	1	6	6	15
ATSIMO ANDREFANA	unknown	0	1	0	2	44	73	120
ATSIMO ANDREFANA	Small	0	4	0	4	43	62	113
ATSIMO ANDREFANA	Medium	0	0	0	1	7	4	12
ATSIMO ANDREFANA	Large	0	0	0	0	0	3	3
BOENI	unknown	0	0	0	0	0	9	9
BOENI	Small	5	14	1	29	76	294	419
BOENI	Medium	2	5	2	6	5	29	49
BOENI	Large	1	1	1	0	3	2	8
DIANA	unknown	1	4	0	4	49	117	175
DIANA	Small	3	4	0	6	33	71	117
DIANA	Medium	0	1	0	1	8	34	44
DIANA	Large	0	1	0	0	2	2	5
Nosy Be	unknown	1	0	0	3	33	30	67
Nosy Be	Small	1	0	0	8	38	28	75
Nosy Be	Medium	0	0	0	0	14	6	20
Nosy Be	Large	0	0	0	0	5	0	5
VAKINAKARATRA	unknown	0	0	0	0	0	1	1
VAKINAKARATRA	Small	9	39	8	60	67	431	614
VAKINAKARATRA	Medium	1	4	1	2	8	28	44
VAKINAKARATRA	Large	0	5	3	1	0	3	12
TOTAL		143	217	216	592	1226	5464	7858

Appendix C Weights
Strict Weights

Region	Employees	Food & Beverages	Raw Material Processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism
Analamanga	Unknown	1.0	5.4	10.5		1.5	
	5-19	2.4	5.6	11.1	44.1	2.3	23.9
	20-99	16.6	1.2	1.6	12.5	1.4	7.0
	100+	6.0	1.0	1.2	3.0	1.6	1.8
Anosy	Unknown	1.0			1.3		1.0
	5-19			1.0	1.0	1.0	1.0
	20-99				1.0		1.0
	100+				1.0		
Antsinanana	Unknown			1.0			
	5-19	2.7	1.1	3.1	35.3		
	20-99		1.0		3.5	1.0	2.4
	100+		1.0		3.0		1.8
Atsimo Andrefana	Unknown	1.0					1.0
	5-19	1.0	1.0	1.0	7.6		1.5
	20-99			1.0	1.0		2.7
	100+				1.9		
Boeny	Unknown						
	5-19	1.0	1.0	1.1	78.8	1.0	
	20-99	1.0	1.0	1.7			
	100+	1.0					1.0
Diana	Unknown	1.0					1.0
	5-19	1.0	1.1	2.1	7.1	1.0	1.0
	20-99	1.0					1.0
	100+				1.0		1.0
Nosy Be	Unknown			1.0	0.0		2.9
	5-19			1.0	2.2	1.0	1.0
	20-99	1.0		1.0	1.0		1.3
	100+						1.7
Vakinankaratra	Unknown						
	5-19	1.1	1.3	1.5	7.6		
	20-99				2.7		2.3
	100+			1.0	1.0	1.0	

Median Weights

Region	Employees	Food & Beverages	Raw Material Processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism
Analamanga	Unknown	1.5	8.3	15.8		2.6	
	5-19	3.2	7.4	14.3	56.4	3.3	29.1
	20-99	24.1	1.7	2.2	17.7	2.2	9.4
	100+	8.9	1.1	1.7	4.4	2.7	2.5
Anosy	Unknown	1.0			1.8		1.0
	5-19			1.0	1.0	1.0	1.0
	20-99				1.1		1.0
	100+				1.0		
Antsinanana	Unknown			1.0			
	5-19	2.6	1.0	2.9	32.8		
	20-99		1.0		3.6	1.0	2.4
	100+		1.0		3.2		1.8
Atsimo Andrefana	Unknown	1.0					1.0
	5-19	1.0	1.0	1.0	7.0		1.3
	20-99			1.0	1.0		2.6
	100+				2.0		
Boeny	Unknown						
	5-19	1.0	1.0	1.2	82.7	1.0	
	20-99	1.0	1.1	1.9			
	100+	1.0					1.0
Diana	Unknown	1.0					1.0
	5-19	1.0	1.2	2.0	6.9	1.0	1.0
	20-99	1.0					1.0
	100+				1.0		1.0
Nosy Be	Unknown			1.0			3.0
	5-19		1.0	1.0	2.0	1.0	1.0
	20-99	1.0		1.0	1.0		1.2
	100+						1.7
Vakinankaratra	Unknown						
	5-19	1.2	1.4	1.7	8.5		
	20-99				3.3		2.7
	100+			1.0	1.1	1.5	

Weak Weights

Region	Employees	Food & Beverages	Raw Material Processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism
Analamanga	Unknown	3.7	22.1	34.4		6.7	
	5-19	4.4	11.3	17.7	85.0	4.8	41.4
	20-99	28.8	2.3	2.4	23.3	2.8	11.7
	100+	9.6	1.3	1.7	5.2	3.1	2.7
Anosy	Unknown	1.0			4.4		1.7
	5-19			1.0	1.0	1.0	1.0
	20-99				1.3		1.0
	100+				1.0		
Antsinanana	Unknown			1.0			
	5-19	4.8	2.2	4.9	67.2		
	20-99		1.0		6.5	1.1	4.0
	100+		1.1		5.1		2.8
Atsimo Andrefana	Unknown	1.0					2.0
	5-19	1.0	1.0	1.0	8.8		1.6
	20-99			1.0	1.0		2.7
	100+				2.0		
Boeny	Unknown						
	5-19	1.0	1.4	1.4	115.4	1.0	
	20-99	1.0	1.4	2.0			
	100+	1.0					1.0
Diana	Unknown	1.8					1.9
	5-19	1.0	2.0	2.8	11.8	1.0	1.0
	20-99	1.0					1.1
	100+				1.0		1.1
Nosy Be	Unknown			2.3			7.7
	5-19		1.0	1.0	3.2	1.0	1.3
	20-99	1.0		1.0	1.0		1.6
	100+						2.0
Vakinankaratra	Unknown						
	5-19	1.9	2.5	2.4	14.3		
	20-99				4.9		3.8
	100+			1.0	1.4	1.9	

Appendix D

Strict Universe Estimates Madagascar

Region	Employees	Food & Beverages	Raw Material processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism	Grand Total
Analamanga	Unknown	8	11	42	0	9	0	**70**
	5-19	22	17	100	573	25	119	**855**
	20-99	17	7	31	100	19	35	**209**
	100+	30	5	11	45	34	7	**131**
	Total	**76**	**40**	**184**	**718**	**88**	**161**	**1266**
Anosy	Unknown	1	0	0	4	0	2	**7**
	5-19	0	0	5	9	1	12	**27**
	20-99	0	0	0	1	0	1	**2**
	100+	0	0	0	1	0	0	**1**
	Total	**1**	**0**	**5**	**15**	**1**	**15**	**37**
Antsinanana	Unknown	0	0	1	0	0	0	**1**
	5-19	5	3	12	141	0	0	**162**
	20-99	0	2	0	21	1	7	**32**
	100+	0	1	0	3	0	4	**8**
	Total	**5**	**6**	**13**	**165**	**1**	**11**	**202**
Atsimo Andrefana	Unknown	1	0	0	0	0	12	**13**
	5-19	14	3	20	23	0	24	**84**
	20-99	0	0	1	3	0	5	**9**
	100+	0	0	0	2	0	0	**2**
	Total	**15**	**3**	**21**	**28**	**0**	**41**	**108**
Boeny	Unknown	0	0	0	0	0	0	**0**
	5-19	9	4	13	79	2	0	**106**
	20-99	3	2	3	0	0	0	**8**
	100+	1	0	0	0	0	2	**3**
	Total	**13**	**6**	**16**	**79**	**2**	**2**	**118**
Diana	Unknown	1	0	0	0	0	12	**13**
	5-19	2	1	2	14	1	19	**39**
	20-99	1	0	0	0	0	5	**6**
	100+	0	0	0	1	0	1	**2**
	Total	**4**	**1**	**2**	**15**	**1**	**37**	**60**
Nosy Be	Unknown	0	0	1	0	0	9	**10**
	5-19	0	2	11	11	1	21	**46**
	20-99	1	0	2	4	0	9	**16**
	100+	0	0	0	0	0	3	**3**
	Total	**1**	**2**	**14**	**15**	**1**	**42**	**75**
Vakinankaratra	Unknown	0	0	0	0	0	0	**0**
	5-19	7	3	14	53	0	0	**77**
	20-99	0	0	0	5	0	2	**8**
	100+	0	0	1	1	1	0	**3**
	Total	**7**	**3**	**15**	**59**	**1**	**2**	**87**
Grand Total		**123**	**61**	**270**	**1094**	**95**	**312**	**1954**

Median Universe Estimates Madagascar

Region	Employees	Food & Beverages	Raw Material Processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism	Grand Total
Analamanga	Unknown	12	17	63	0	16	0	**108**
	5-19	29	22	128	734	36	146	**1095**
	20-99	24	10	45	142	31	47	**299**
	100+	45	5	16	66	57	10	**198**
	Total	**110**	**55**	**252**	**941**	**139**	**202**	**1699**
Anosy	Unknown	1	0	0	5	0	2	**8**
	5-19	0	0	5	9	1	12	**27**
	20-99	0	0	0	1	0	1	**2**
	100+	0	0	0	1	0	0	**1**
	Total	**1**	**0**	**5**	**16**	**1**	**15**	**38**
Antsinanana	Unknown	0	0	1	0	0	0	**1**
	5-19	5	3	12	131	0	0	**151**
	20-99	0	2	0	22	1	7	**32**
	100+	0	1	0	3	0	4	**8**
	Total	**5**	**6**	**13**	**156**	**1**	**11**	**192**
Atsimo Andrefana	Unknown	1	0	0	0	0	12	**13**
	5-19	14	3	20	21	0	21	**79**
	20-99	0	0	1	3	0	5	**9**
	100+	0	0	0	2	0	0	**2**
	Total	**15**	**3**	**21**	**26**	**0**	**38**	**103**
Boeny	Unknown	0	0	0	0	0	0	**0**
	5-19	9	4	13	83	2	0	**111**
	20-99	3	2	4	0	0	0	**9**
	100+	1	0	0	0	0	2	**3**
	Total	**13**	**6**	**17**	**83**	**2**	**2**	**123**
Diana	Unknown	1	0	0	0	0	12	**13**
	5-19	2	1	2	14	1	19	**39**
	20-99	1	0	0	0	0	5	**6**
	100+	0	0	0	1	0	1	**2**
	Total	**4**	**1**	**2**	**15**	**1**	**37**	**60**
Nosy Be	Unknown	0	0	1	0	0	9	**10**
	5-19	0	2	11	10	1	21	**45**
	20-99	1	0	2	4	0	9	**16**
	100+	0	0	0	0	0	3	**3**
	Total	**1**	**2**	**14**	**14**	**1**	**42**	**74**
Vakinankaratra	Unknown	0	0	0	0	0	0	**0**
	5-19	8	3	15	59	0	0	**86**
	20-99	0	0	0	7	0	3	**9**
	100+	0	0	1	1	1	0	**4**
	Total	**8**	**3**	**16**	**67**	**1**	**3**	**99**
Grand Total		**157**	**76**	**340**	**1318**	**146**	**350**	**2388**

Weak Universe Estimates

Region	Employees	Food & Beverages	Raw Material Processing Manufacturing	Other Manufacturing	Other Services	Textiles & Garments	Tourism	Grand Total
Analamanga	Unknown	30	44	138	0	40	0	**252**
	5-19	39	34	160	1105	53	207	**1598**
	20-99	29	14	49	186	39	58	**375**
	100+	48	6	15	78	65	11	**224**
	Total	**146**	**98**	**361**	**1369**	**198**	**276**	**2449**
Anosy	Unknown	1	0	0	13	0	3	**17**
	5-19	0	0	5	9	1	12	**27**
	20-99	0	0	0	1	0	1	**2**
	100+	0	0	0	1	0	0	**1**
	Total	**1**	**0**	**5**	**24**	**1**	**16**	**48**
Antsinanana	Unknown	0	0	1	0	0	0	**1**
	5-19	10	6	19	269	0	0	**304**
	20-99	0	2	0	39	1	12	**54**
	100+	0	1	0	5	0	6	**12**
	Total	**10**	**10**	**20**	**313**	**1**	**18**	**371**
Atsimo Andrefana	Unknown	1	0	0	0	0	24	**25**
	5-19	14	3	20	26	0	25	**89**
	20-99	0	0	1	3	0	5	**9**
	100+	0	0	0	2	0	0	**2**
	Total	**15**	**3**	**21**	**31**	**0**	**55**	**125**
Boeny	Unknown	0	0	0	0	0	0	**0**
	5-19	9	5	15	115	2	0	**147**
	20-99	3	3	4	0	0	0	**10**
	100+	1	0	0	0	0	2	**3**
	Total	**13**	**8**	**19**	**115**	**2**	**2**	**160**
Diana	Unknown	2	0	0	0	0	23	**25**
	5-19	2	2	3	24	1	19	**50**
	20-99	1	0	0	0	0	5	**6**
	100+	0	0	0	1	0	1	**2**
	Total	**5**	**2**	**3**	**25**	**1**	**49**	**84**
Nosy Be	Unknown	0	0	2	0	0	23	**25**
	5-19	0	2	11	16	1	27	**57**
	20-99	1	0	2	4	0	11	**18**
	100+	0	0	0	0	0	4	**4**
	Total	**1**	**2**	**15**	**20**	**1**	**65**	**104**
Vakinankaratra	Unknown	0	0	0	0	0	0	**0**
	5-19	13	5	21	100	0	0	**139**
	20-99	0	0	0	10	0	4	**14**
	100+	0	0	1	1	2	0	**4**
	Total	**13**	**5**	**22**	**111**	**2**	**4**	**157**
Grand Total		**203**	**128**	**467**	**2009**	**206**	**484**	**3498**

Appendix E

Original Sample Design, Madagascar:
   Note significant numbers of firms with unknown size were selected to ensure that the sample was not biased.

REGION	Size	Raw material processing manufacturing	Food & Beverages	Textiles & Garments	Other Manufacturing	Hospitality &'Tourism	Other services	TOTAL
ANALAMANGA	Unassigned	5	4	6	2	0	0	17
ANALAMANGA	Small	5	3	10	6	4	6	34
ANALAMANGA	Medium	20	8	21	13	3	5	70
ANALAMANGA	Large	8	9	19	12	3	7	58
ANOSY	Unassigned	1	2	0	2	8	9	22
ANOSY	Small	0	3	0	2	12	8	25
ANOSY	Medium	0	0	0	0	1	2	3
ANOSY	Large	1	0	0	0	0	2	3
ANTSINANANA	Unassigned	0	0	0	1	0	0	1
ANTSINANANA	Small	9	7	2	5	0	5	28
ANTSINANANA	Medium	2	3	0	6	6	4	21
ANTSINANANA	Large	2	1	0	1	1	4	9
ATSIMO ANDREFANA	Unassigned	3	1	0	2	8	0	14
ATSIMO ANDREFANA	Small	8	4	0	4	7	0	23
ATSIMO ANDREFANA	Medium	0	0	0	1	7	4	12
ATSIMO ANDREFANA	Large	0	0	0	0	0	1	1
BOENI	Unassigned	0	0	0	0	0	0	0
BOENI	Small	10	8	1	6	0	0	25
BOENI	Medium	4	5	2	6	5	0	22
BOENI	Large	1	1	1	0	2	0	5
DIANA	Unassigned	7	4	0	4	17	0	32
DIANA	Small	4	4	0	6	12	4	30
DIANA	Medium	1	1	0	1	8	1	12
DIANA	Large	0	1	0	0	2	1	4
Nosy Be	Unassigned	1	0	0	3	16	0	20
Nosy Be	Small	1	0	0	7	11	2	21
Nosy Be	Medium	0	0	0	0	14	5	19
Nosy Be	Large	0	0	0	0	4	0	4
VAKINAKARATRA	Unassigned	0	0	0	0	0	0	0
VAKINAKARATRA	Small	6	4	8	2	0	5	25
VAKINAKARATRA	Medium	4	4	1	2	3	3	17
VAKINAKARATRA	Large	1	5	3	1	0	2	12
TOTAL		104	82	74	95	154	80	589

Appendix F

Local Agency team involved in the study:

Local Agency	**Name:** Hermes Conseils **Country:** Madagascar **Informations générales** Nom du Cabinet: Cabinet HERMES CONSEILS Contact: Maryse RAMILISON, Gérante Adresse: Lot VE 71 Bis Ambatovinaky Mail: contact@hermesconseils.mg Tél.: 22.259.90 / 22.257.67 Fax: 22.257.67 Mobile: 032.07.821.93 / 034.02. 257.67 Site web: www.hermes-conseils.com **Activities since:** 2000
Enumerators involved:	**Enumerators:** ~32 **Recruiters**: 16
Other staff involved:	**Fieldwork Coordinators:** 8 **Data Entry:** 8 **Data Processing:** 1

Sample Frame:

Characteristic of sample frame used:	The sample frame came from Instat
Source:	Instat list
Year:	2012
Additional list	Lists from chambers of commerce were used for additional contacts and to confirm contact information.

Sectors included in the Sample:

Original Sectors	The manufacturing sector comprises all manufacturing establishments as mentioned in group D The service sector includes Group F (construction), Groups G, Group H (hotels and restaurants), Group I (transport, storage, and communications) and subsector 72 from Group K *As mentioned above, additional sectors from Group A and Group C of ISIC revision 3.1 were included in this survey. The observations obtained from these observations are NOT used for computing indicators on the www.enterprisesurveys.org website and are not covered by this report.*
Added (top up) Sectors	None

Fieldwork and country situation:

Date of Fieldwork	August 2013 to May 2014
Country	Madagascar
Use of CAPI	• NO
Problems found during fieldwork:	▪ Many interviews were completed in more than one visit. ▪ The political climate in Madagascar at the time of the survey resulted in reluctance of many businesses to participate fully. Many respondents and potential respondents were concerned about an ongoing crack-down and compliance enforcement by tax officials. ▪ Respondents were in general inclined to think the study aimed to get tax and social security defaulters. This might have affected the reliability of answers related to revenue and costs. ▪ A unique arrangement whereby the World Bank only provided technical assistance rather than full project management and control resulted in communication and chain-of-command conflicts that slowed implementation and affected overall survey management.
Country specific situation	▪ Elections in Madagascar during the fieldwork period contributed to response-rate issues and may have influenced responses to some questions.