

Madagascar Enterprise Surveys Data Set

1. Introduction

This document provides additional information on the data collected in Madagascar from 15 September 2008 to 13 February 2009 as part of the Enterprise Survey, an initiative of the World Bank.

The objective of the Enterprise Surveys is to obtain feedback from enterprises in client countries on the state of the private sector as well as 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.

Through interviews with firms in the manufacturing and services sectors, the Enterprise Survey data provides information on the constraints to private sector growth and is used to create statistically significant business environment indicators that are comparable across countries.

The report describes the sampling design of the survey, the structure of the dataset and additional information that may be useful when using the data, including information on non-response rates, the calculation of sample weights and country-specific factors that may have affected survey implementation.

2. Survey Target Population

The whole population, or the universe, covered in the Enterprise Surveys is the non-agricultural economy. It comprises: all manufacturing sectors according to the ISIC Revision 3.1 group classification (group D), construction sector (group F), services sector (groups G and H), and transport, storage, and communications sector (group I). Note that this population 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.

For Madagascar, the sectors included in the sample by two-digit ISIC code are as follows:

Manufacturing: 15, 18
Other Manufacturing: 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37
Services: 52
Residual/ Remainder/ Other Services: 45, 50, 51, 55, 60, 61, 62, 63, 64, 72

The Enterprise Survey for Madagascar targeted 480 registered establishments with at least five employees, 120 registered establishments with one to four employees (micro), and 120 non-registered firms in the sectors defined above. The sampling for registered and non-registered establishments was implemented separately.

In the Enterprise Surveys, the requirements for registration are defined on a country-by-country basis using the information collected by Doing Business and information from the in-country

contractors. In Madagascar, registered firms were defined as firms that had a Tax Identification Card (carte d'identification fiscale).

3. Sampling for Registered Establishments

The sample for registered establishments in Madagascar was selected using stratified random sampling, following the methodology explained in the *Sampling Manual*.¹ As discussed in greater detail in the *Sampling Manual*, stratified random sampling was preferred over simple random sampling in the Enterprise Surveys for several reasons:^{2,3}

- a. To obtain unbiased estimates for different subpopulations within the economy with some known level of precision.
- b. To obtain unbiased estimates for the whole population.
- c. To ensure that the final sample includes establishments from all relevant sectors in the country 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 many cases, will be more precise than using a simple random sampling method (i.e., lower standard errors, other things being equal.)
- e. Stratification may also produce a smaller bound on the estimation errors 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 of collecting the survey data may be reduced by stratifying the population elements into convenient groupings.

Three levels of stratification were used in the Madagascar sample: firm sector, firm size, and geographic region. The original sample design, with specific targets for these strata, is included in Appendix A.

Industry stratification was designed as follows: the universe was stratified into three manufacturing industries (food, textiles, and other), one services industry (retail) and one residual sector as defined in the sampling manual. The initial sample design had a target of 240 interviews in manufacturing and 120 interviews each in the services and residual categories, though this sample design was later adjusted to reflect the low prevalence of manufacturing establishments in Madagascar. The initial sample design for micro targeted 120 establishments, 60 in manufacturing and 60 in services. The sample design for the informal survey was designed to mirror the micro sample.

Size stratification was defined following the standardized definition used for the Enterprise Surveys: micro (1 to 4 employees), small (5 to 19 employees), medium (20 to 99 employees),

¹ The complete text of the *Sampling Manual* 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 See Cochran, W., 1977, pp. 89; Lohr, Sharon, 1999, pp. 95.

and large (more than 99 employees). For stratification purposes, the number of employees was defined on the basis of reported permanent full-time workers.

Regional stratification was defined in terms of the geographic regions with the largest commercial presence in the country: Antananarivo, Mahajanga, Toamasina, and Antsiranana were the four metropolitan areas selected in Madagascar.

4. Sampling implementation

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

Two frames were used for Madagascar. The first was an extract from the database of active establishments provided by Institute National de la Statistique du Madagascar [INSTAT]. The second frame (the panel sample) consisted of enterprises interviewed for the Enterprise Survey in 2005, which were to be re-interviewed where they were in the selected geographical regions and met eligibility criteria. Counts from both sample frames shown below.

Universe Figures for Madagascar

Source: Institute National de la Statistique du Madagascar [INSTAT]

Region	Employees	Sector					Grand Total
		15	18	Other Manufacturing	52	Other Services	
Antananarivo	0-4	67	58	311	785	977	2,198
	5-19	55	47	239	408	525	1,274
	20-99	27	54	136	90	233	540
	100+	18	113	63	23	91	308
Antananarivo Total		167	272	749	1,306	1,826	4,320
Mahajanga	0-4	9	6	38	114	85	252
	5-19	7	6	26	42	46	127
	20-99	4	1	9	4	26	44
	100+	1	1	3		11	16
Mahajanga Total		21	14	76	160	168	439
Toamasina	0-4	13	1	31	150	111	306
	5-19	6		27	61	87	181
	20-99	4	1	17	17	32	71
	100+			7	5	14	26
Toamasina Total		23	2	82	233	244	584
Antsiranana	0-4	4	2	18	83	53	160
	5-19	1		13	28	37	79
	20-99	3		5	2	8	18
	100+	1			1	6	8
Antsiranana Total		9	2	36	114	104	265
Grand Total		220	290	943	1,813	2,342	5,608

Panel sample counts for Madagascar

Region	Employees	Sector				Grand Total
		15	18	Other Manufacturing	52	
Antananarivo	0-4					
	5-19	8	5	70		83
	20-99	15	10	66		91
	100+	7	37	34		78
Antananarivo Total		30	52	170		252
Mahajanga	0-4					
	5-19	2		1		3
	20-99	4		6	1	11
	100+					
Mahajanga Total		6		7	1	14
Toamasina	0-4					
	5-19			3		3
	20-99			6		6
	100+	1				1
Toamasina Total		1		9		10
Antsiranana	0-4					
	5-19	1		3		4
	20-99	1		1		2
	100+	1				1
Antsiranana Total		3		4		7
Grand Total		40	52	190	1	283

The enumerated establishments were then used as the frame for the selection of a sample with the aim of obtaining interviews at 480 establishments with five or more employees and a further 120 with establishments with less than five employees (micro establishments).

The quality of the frame was assessed at the onset of the project and was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc. 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 26.3% (558 out of 2,120 establishments for the ES and micro samples, including panel establishments).⁴

For the informal sector, there were no sample lists of firms. The sampling procedure was to survey a non-registered establishment similar to each registered establishment surveyed in the micro establishment. Because a formal sample frame was not used, it is not possible to calculate response rates, universe estimates, or sampling weights for the informal sector sample.

5. Database Structure

The database compiles four different versions of the Enterprise Survey questionnaire for all registered establishments (including micro). The basic questionnaire, the Core Module, includes all common questions asked to all establishments from all sectors (manufacturing, retail and residual) and this module is administered to establishments in the residual sector. Micro firms completed a module very similar to the Core Module. The Manufacturing Questionnaire is built

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

upon the Core Module and adds some specific questions relevant to the sector. The Services Questionnaire is also built upon the Core Module and adds specific questions relevant to retail. Each version of the questionnaire is identified by the index variable, *a0*. The questionnaire for the informal sector is compiled separately.

All variables in the database 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 “AF” indicate questions specific to the Africa region; these questions may not have been asked in Enterprise Surveys conducted in countries in other regions. Variable names preceded by a prefix “MA” indicate questions specific to the Madagascar surveys; these questions may not have been asked in Enterprise Surveys conducted in other countries. All other variables are global and are present in all country surveys conducted throughout the world. All variables are numeric with the exception of those variables with an “x” at the end of their names. The suffix “x” indicates that the variable is alpha-numeric.

The variable *idstd* uniquely identifies each establishment at the global level.

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. These variables generate the strata cells for each industry/region/size combination. 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 the sampling regions

- a6a*: coded using the definition for micro, small, medium, and large establishments as discussed 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 industries that comprise the manufacturing, services, and residual categories used in the stratification. These codes include most manufacturing industries (15 to 37), and retail, and IT for services (52, and 72 respectively). All establishments within the ‘other manufacturing’ stratum were coded with *a4a*=2.

- panel2005*: identifies panel firms

- id*: for panel firms the variable contains the same id used in 2005

Note that these variables may not coincide with reality for some establishments as sample frames may contain information that is later found to be inaccurate.

The surveys were implemented following a two stage procedure. In the first stage a screener questionnaire was administered over the phone to determine sampled establishment’s eligibility for the survey and to make appointments; in the second stage, a face-to-face interview took place with the Manager/Owner/Director of each establishment. The variables *a4b* and *a6b* contain the industry and size of the establishment from the screener questionnaire. Variables *a8* to *a11* contain additional information that was collected in the screening phase.

The main questionnaire contains variables for location (*a3x*), industry (*d1a2*), and number of employees (*l1*, *l6* and *l8*) that more accurately reflect describe the characteristics of establishments than the information provided on these variables in the sample frame or the screener.

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 industry-strata based on information available from the sample frame, whereas variable *d1a2* indicates the actual ISIC code of the main output of the establishment as answered by the interviewee. This variable is probably the most accurate variable with which to classify establishments by activity.

Variable *a3x* indicates the actual location of the establishment. There may be divergences between the location in the sampling frame and the actual location, as establishments may be listed in one place on the sample frame but the actual physical location is in another place.

Variables *l1*, *l6* and *l8* provide a more accurate measure of employment and account for both permanent and temporary employment. Special efforts were made to make sure that this information was not missing for most establishments.

6. Universe Estimates

Special care is given to the correct computation of universe estimates and weights in the Enterprise Surveys. Considering the varying quality of sample frames across countries, it is important to accurately adjust the universe totals within each region/industry/size stratum to account for the presence of ineligible units in the sampling frame.⁵ The information collected during the screening process is used to scale down the universe estimate for each cell by the observed proportion of ineligible units within the cell.

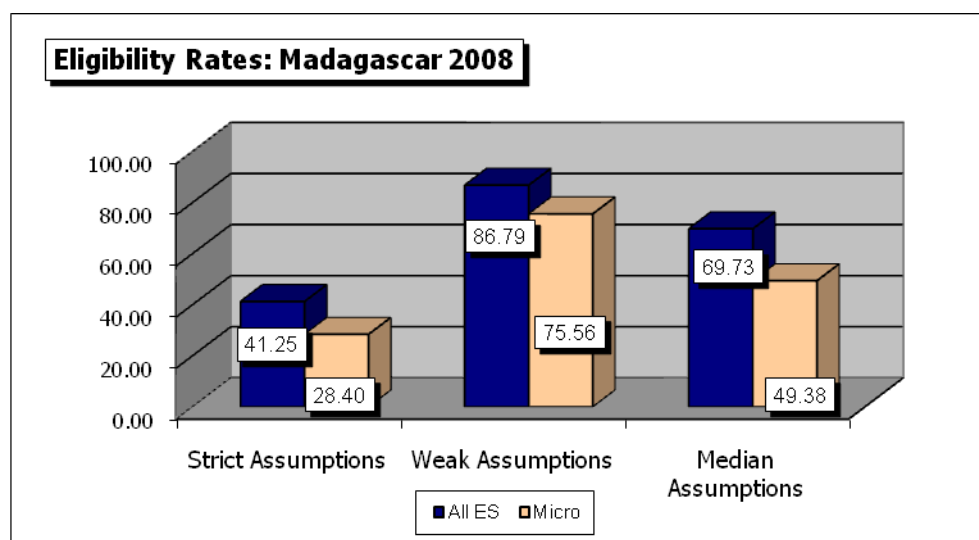
Of course, different assumptions about the eligibility of establishments result in different adjustments to the universe cells and thus different sampling weights. For some establishments where contact was not successfully completed during the screening process it is not possible to directly determine eligibility. Three sets of assumptions on establishment eligibility are considered:

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

b- 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*.

⁵ For example, ineligible units could include: firms that discontinued businesses, education or government establishments, establishments with less than 5 employees in the ES sample, establishments where there was no reply after having called in different days of the week and in different business hours, the number was out of order, no tone in the phone line, answering machine, fax line, wrong address or moved away and could not obtain new contact information.

c- 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. The resulting weights are included in the variable *w_weak*. Under the weak assumption only observed non-eligible units are excluded from universe projections. The following graph shows the different eligibility rates under each set of assumptions.



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 C shows the universe estimates of the numbers of registered establishments based on the strict, weak and median relative estimates.

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.

7. Weights

Since the sampling design was stratified and employed differential sampling of the strata, 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 *pa* in Stata.)⁶

⁶ This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

Two sets of weights for each cell were computed using the strict, weak, and median assumptions on establishment eligibility. The first set of estimates calculated proportions using the raw sample count for each cell. However, for many cells the sample numbers of interviewed establishments are small, and eligibility rates and adjusted universe cells projections for those cells are subject to relatively large sampling variations. Therefore a second set of more robust estimates (collapsed weights) that use the multiples of the relative eligibility rates for each industry, size, and region was also produced. The collapsed weights are based on larger samples than the individual cells and thus produce values with smaller sampling variations. The data sets include only the robust weights.

Appendix D shows the cell weights for registered establishments in Madagascar.

8. Appropriate use of weights

As discussed above, 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.

However, there is some discussion on the proper 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 use of weighted OLS for a common population coefficient).

From a more general approach, if the regressions are descriptive of the population then weights should be used. The estimated model can be thought of as the relationship that would be expected if the whole population were observed.⁷ If the models are developed as structural relationships or behavioral models that may vary for different parts of the population, there is no reason to use weights.

9. Non-response

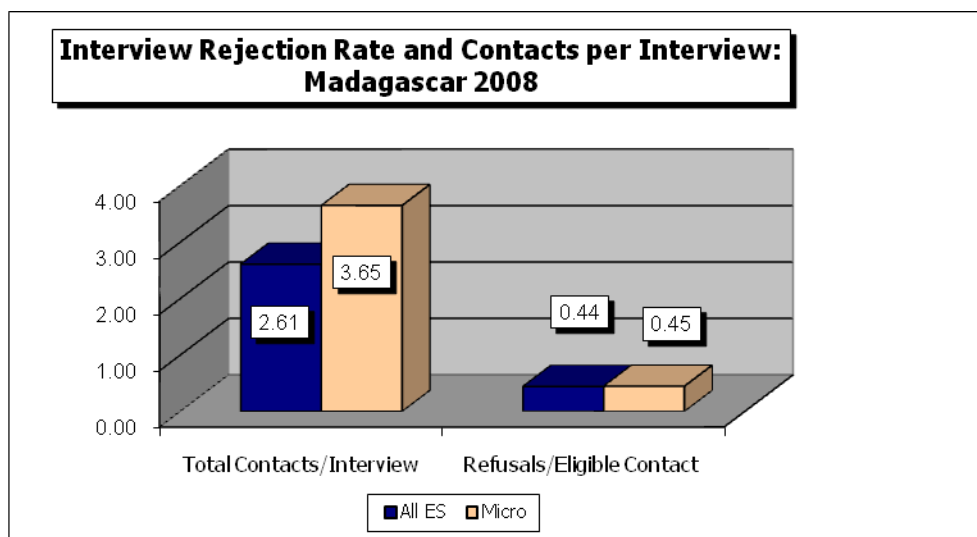
The Enterprise Surveys, along with all other surveys, suffer from both survey non-response and item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the refusals to answer some specific questions. Different strategies were used to address these issues.

⁷ The use weights in most model-assisted estimations using survey data is strongly recommended by the statisticians specialized on survey methodology of the JPSM of the University of Michigan and the University of Maryland.

Survey non-response was addressed by maximizing efforts to contact establishments that were initially sampled. When the survey frame was extracted from the sampling frame, several establishments with the same strata characteristics were randomly selected for each interview and each establishment was assigned a preference number.⁸ Substitutions of replacement establishments were made in order to help achieve targets on the number of interviews for each stratum. Extensive efforts were made to complete interviews with each first preference establishment before contact with a replacement establishment was allowed. At least four attempts were made to contact each sampled establishment for an interview at different times/days of the week before a replacement establishment was allowed to be contacted for an interview.

Further research is needed on survey non-response in the Enterprise Surveys regarding the potential introduction of bias through substitution and non-response.

As the following graph shows, the number of contacted establishments per realized interview was 2.61 for the Enterprise Survey and panel sample and 3.65 for micro firms. 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 (e.g., establishments that closed or were in ineligible sectors). Refusal rates are also shown in the graph below. For each establishment eligible for an interview, 0.44 in the Enterprise Survey sample and 0.45 in the micro sample refused to participate.

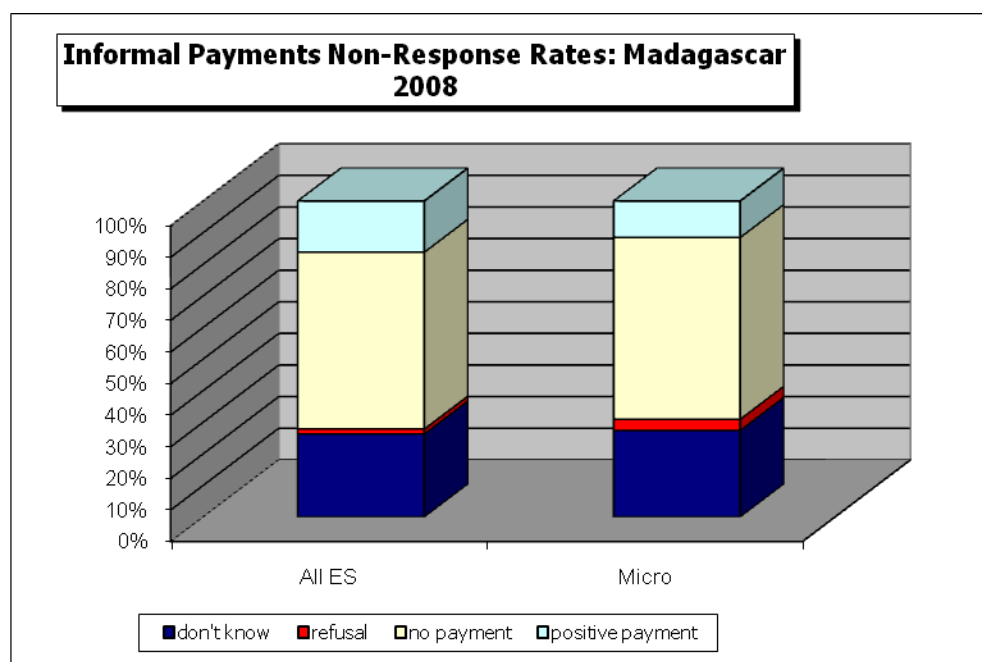


⁸ In cases where the number of contacts initially drawn from the sample frame are insufficient to obtain an interview with the targeted number of establishments in a given strata, additional contacts for that strata may be drawn from the sampling frame. If all establishments in that strata have already been contacted and the sample target has not been reached, the sample design may be adjusted to allow additional interviews in other strata. In Madagascar, the sample design was readjusted from the original design to reflect the small numbers of manufacturing establishments in the country. Additionally, more contacts were issued near the end of fieldwork to compensate for lower than expected response rates in several strata.

In completed surveys, 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 (-7) as a different option from don't know (-9).

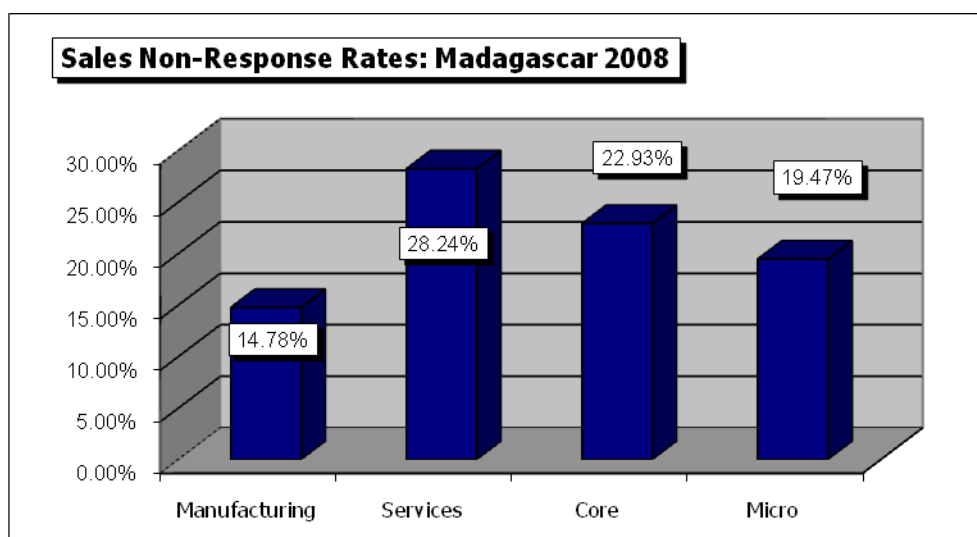
The following graph shows the breakdown of answers about the total amount of informal payments made annually (variable *j7b*) by questionnaire type.



b- Establishments with incomplete information on important questions including total sales, cost figures and employment levels were re-contacted in order to complete this information. However, re-contacts did not fully eliminate low response rates for some items.

The following graph shows non-response rates for the sales variable, *d2*, by type of questionnaire.⁹

⁹ Please note that the question on total sales does not have a “refuse to answer” option, thus the non-response rates in the graph above reflect DKs and NAs as well as any missing values.



This report summarizes statistics on rejection rates, eligibility rates, and item non-response to alert researchers of these issues when using the data and when making inferences. Item non-response, selection bias, and imperfect sampling frames are not unique to Madagascar or the Enterprise Surveys. All surveys suffer from these issues although they may not be made explicit.

10. Country specific comments

Local Agency team involved in the study:

Local Agency	Name: ATW CONSULTANTS Country: MADAGASCAR Membership of international organization: ESOMAR Activities since: 1991
Name of Project Manager	Andrianina RAKOTOARIMANANA – General Manager
Name and position of other key persons of the project:	Fanja RAJAONA – Research Executive
Enumerators involved:	Enumerators/ recruiters: 50
Other staff involved:	Fieldwork Coordinators: 4 Editing: Data Entry: 5 Data Processing: 2

Comments on sample frame:

Characteristic of sample frame used:	<p>The database for the fresh sample was bought from Institute National de la Statistique du Madagascar [INSTAT] and contained the following characteristics:</p> <ul style="list-style-type: none"> – Name of the firm – Contact details – ISIC code – Number of employees <p>The same characteristics were contained in the World Bank 2005 database used for the Panel Sample.</p>
Source:	<p>The sample frame source for the fresh Sample is the database from the Institute National de la Statistique du Madagascar [INSTAT], that registers all formal firms, detailing the activity sector and the number of employees. This database is yearly updated.</p> <p>The World Bank 2005 database was used to interview panel firms</p>
Year:	2008 for Fresh Sample, 2005 for Panel
Comments on the quality of sample frame:	<p>The database source provides exact number of employees, but this number was certainly underestimated since the establishments usually do not declare all their employees to pay less taxes.</p> <p>Finally, it appeared that the list was not so accurate as a lot of firms listed in the sample frame discontinued business or had wrong contacts details.</p>
Year and organization that conducted the last economic census:	National Statistic Institute – INSTAT – 2008, but this database does not take into account the firms that have closed. In some cases households were included. Many time the employee number was missing.
Other sources for company statistics:	None.

Comments on sample:

Comments/ problems on sectors and regions selected in the sample:	<p>On sectors: a lot of firms classified in the sample frame as manufacturing (especially food manufacturing 15, and textile manufacturing 18) finally appeared on the field to be working in services. We then run out of list several times for eligible contacts in these sectors</p> <p>On regions: None.</p>
Comments on the response rate:	<p>From experience, we can say that the response rate for this survey is rather good, taking into account,</p> <ol style="list-style-type: none"> 1) the target (business target, and especially top-managers, that are difficult to reach), 2) the nature of information being asked in this survey (in Madagascar, financial data are usually very confidential) 3) the duration of the questionnaire (about 1 hour) 4) a context in which local firms are not very familiar with surveys

Comments on the sample design:	The sample design had to be readjusted many times, since the available contacts in the first selected sectors proved to be insufficient to complete the target sample. Target number of interviews per sector and region were adjusted according to availability on the Universe.
Other comments:	None.

Comments on Fieldwork:

Date of Fieldwork	15 th September 2008 to 13 th February 2009
Country	MADAGASCAR
Interview number	<p>Main ES Survey + Micro Enterprises</p> <p>Sector 15: 64 (46 ES + 18 Micro)</p> <p>Sector 18: 65 (59 ES + 6 Micro)</p> <p>Other Manufactures: 139 (112 ES + 27 Micro)</p> <p>Sector 52: 149 (114 ES + 35 Micro)</p> <p>Other Services: 141 (116 ES + 25 Micro)</p> <p>Informal: 124</p> <p>Employees: 1093</p>
Problems found during fieldwork:	<ul style="list-style-type: none"> - Phone numbers were very difficult to find. We were then obliged to find firms on the field and to administrate the screener face to face. But even addresses in the list were many times wrong/ very hard to find, it then took a lot of time to locate firms. - Transport costs in Madagascar are very high and time consuming (the regions very far from one another, very expensive flights, trip between one town to another taking at least one day long, etc.). We had then to do each region in block, with no possibility to come back to one region after it had been achieved. This made sample readjustment on one region after completion impossible. - It was very difficult to collect data related to the establishment financial statement, because 1) firms are reluctant to give them (it's the local culture, even big firms do not publish their statements as they should do), 2) figures asked in the questionnaire couldn't be found directly in the statement but must have been calculated, 3) a large proportion of firms do not keep accountancy - There is a high scepticism as to the ultimate value of the research results to participating companies: the World Bank is known to do several surveys which impact is not clearly seen - We had a high turnover of our interviewers' team, since they got easily exhausted because of the survey difficulty.
Other observations:	<p>- A part of the fieldwork coincided with the Christmas holidays. During this period, the majority of firms were closed, and in firms that were still open, the top -managers were not available. We then had to stop the fieldwork between the 15th December until the 9th January.</p> <p>- From the 26th January, a big crisis began in Antananarivo: the town felt under curfew, there were manifestations in the</p>

	city in which many people were injured and even killed. For security reason, fieldwork had to stop again. The political situation isn't stable yet in March 2009. Therefore, it was decided to stop the fieldwork the 13 th February even if the target wasn't reached yet.
--	--

Questionnaires:

Problems for the understanding of questions	<p>- D16 (days of inventory, measured in days of production): this notion could be understood only by very large firm</p> <p>- F1 (capacity utilization): this notion was very difficult to understand for small firms</p> <p>- J1: The two last sentences ("Government officials' interpretations of the laws being consistent and predictable" and "confidence in the legal system") are misunderstood most of the time, we had to repeat the sentences many times, and after that the respondent still didn't catch their meaning. Yet, they give an answer, without understanding the real meaning of the question</p> <p>- N6 (net book value): this notion could be understood only by very large firms</p> <p>Apart from this, some questions were understood but very difficult to answer, because most of the firms didn't have a clear idea about the answer:</p> <p>-C9 (losses resulted from power outages): majority of firms cannot assess this loss</p> <p>-J2 (time spent on dealing with requirements imposed by government regulations): majority of respondents had difficulty to assess this time and gave a figure at random</p> <p>-L4, L5, L6, etc (split of skilled/ unskilled employees, permanent full-time production/ non production employees, seasonal/temporary workers...): only few firms had this kind of information directly available in their books</p> <p>-L9 (average educational attainment of a typical production worker): respondents usually are not able to answer directly in number of years, but give an average class (such as "3rd grade class")</p> <p>56</p> <p>-MAL17 (total compensation for each category) : this is very difficult to calculate because firms do not have this kind of figures</p> <p>-AFn2k (depreciation): a lot of firms do not apply depreciation or even do not know how to calculate it</p> <p>We would say that generally speaking, the questionnaire in his whole is adapted for large firms, but less relevant as far as little firms are concerned, and even more to informal firms</p>
Problems found in the navigability of – questionnaires (for example, skip patterns).	None.
Comments on	Questionnaire length has considerably been shortened

questionnaires length:	compared to the one used in 2005. Nevertheless, it's still very time demanding for top managers who are always very busy.
Suggestions or other comments on the questionnaire:	<ul style="list-style-type: none"> - An introduction part, to be read by the interviewer should be added at the beginning. Making it easier for the interviewer to introduce the survey and making it homogenous for all interviewers - Questionnaire layout still should be improved to make the questionnaire look shorter (thinner): top managers are discouraged just at view of the size of the questionnaire.

Database:

Comments on the data map	<p>Data entry program chosen: PERTS</p> <p>Comments:</p> <p>Globally, the program is not very practical to use:</p> <ul style="list-style-type: none"> - It took a lot of time to understand how to use correctly the data entry program and make all skips work correctly - Correction of the value of only one variable required the validation of all variables from the one modified until the end of the questionnaire; so with the length of the questionnaire modifications were very time demanding - The program didn't allow us to delete questionnaires - The program didn't allow us to check inconsistency as we usually do. - Some scales in the program are different from the scales available in the questionnaire (questions about obstacles in each part), and couldn't be changed: then the operators have to think about the scale equivalent while punching, which may cause some mistakes
Non response rate	<ul style="list-style-type: none"> - When the questionnaire appeared to have too many "non response", our supervisors asked for additional appointment to try to collect more data. Nevertheless, we are still a context in which firms are very reluctant to provide financial data
Comments on the data cleaning	<ul style="list-style-type: none"> - The Data Validation Report focused especially on Don't Know and Refuse answer. But as we said below, in some cases we cannot avoid this kind of answers. - Short time for data cleaning

Country situation:

General aspects of economic, political or social situation of the country that could affect the results of the survey:	None.
Relevant country	Trouble in the country by the end of January, considerably

events occurred during fieldwork:	affecting the fieldwork.
Other aspects:	None.

References

Cochran, William G., Sampling Techniques, 1977.

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

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

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

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

Appendix A

Original Sample Design

Region name	Employee size	15	18	52	Other Manufacturing	Other Services	Grand Total
ANTANANARIVO	5-19	27	23	36	14	22	122
	20-99	13	27	42	16	28	126
	100+	9	27	11	22	30	99
ANTANANARIVO Total		49	77	89	52	80	347
ANTSIRANANA	5-19			6	4	5	15
	20-99	1		1	2	2	6
	100+					2	2
ANTSIRANANA Total		1	0	7	6	9	23
MAHAJANGA	5-19	3	3	6	5	5	22
	20-99	2		2	3	6	13
	100+				1	4	5
MAHAJANGA Total		5	3	8	9	15	40
TOAMASINA	5-19	3		6	5	5	19
	20-99	2		8	5	7	22
	100+			2	3	4	9
TOAMASINA Total		5	0	16	13	16	50
Grand Total		60	80	120	80	120	460

Region name	Employee size	15	18	52	Other Manufacturing	Other Services	Grand Total
ANTANANARIVO	1-4	5	5	15	8	19	52
ANTSIRANANA	1-4	1		8	3	6	18
MAHAJANGA	1-4	2	2	9	5	7	25
TOAMASINA	1-4	3		10	5	7	25
Grand Total		11	7	42	21	39	120

Appendix B

Enterprise Survey Sample

Status Codes

ELIGIBLES		
Eligible	1. Eligible establishment (Correct name and address)	319
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	0
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	6
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	2
	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	17
	6. The firm discontinued businesses	76
	7. Not a business: private household	1
	8. Ineligible activity: education, agriculture, finances, governments...	15
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	0
	92. Line out of order	0
	93. No tone	0
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	170
	13. Refuses to answer the screener	268
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	22
	151. Out of target - outside the covered regions, firm moved abroad	5
	152. Out of target - firm moved abroad	1
	Total	902

Response Outcomes

Complete interviews (Total)	306
Incomplete interviews	4
Eligible in process	1
Refusals	16
Out of target	109
Impossible to contact	170
Ineligible - coop.	6
Refusal to the Screener	268
Total	880

Panel Sample

Status Codes

ELIGIBLES		
Eligible	1. Eligible establishment (Correct name and address)	148
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	0
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	2
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	0
	16. Panel firm - now less than five employees	4
Ineligible	5. The establishment has less than 5 permanent full time employees	0
	6. The firm discontinued businesses	34
	7. Not a business: private household	0
	8. Ineligible activity: education, agriculture, finances, governments...	1
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	0
	92. Line out of order	0
	93. No tone	0
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	29
	13. Refuses to answer the screener	64
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	0
	151. Out of target - outside the covered regions, firm moved abroad	0
	152. Out of target - firm moved abroad	0
Total		282

Response Outcomes

Complete interviews (Total)	141
Incomplete interviews	0
Eligible in process	2
Refusals	11
Out of target	35
Impossible to contact	29
Ineligible - coop.	0
Refusal to the Screener	64
Total	282

Micro Sample

Status Codes

ELIGIBLES		
Eligible	1. Eligible establishment (Correct name and address)	112
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	0
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	3
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	0
	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	28
	6. The firm discontinued businesses	63
	7. Not a business: private household	3
	8. Ineligible activity: education, agriculture, finances, governments...	1
Unobtainable	91. No reply (after having called in different days of the week and in different business hours)	0
	92. Line out of order	0
	93. No tone	0
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	106
	13. Refuses to answer the screener	85
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	4
	151. Out of target - outside the covered regions, firm moved abroad	2
	152. Out of target - firm moved abroad	2
Total		409

Response Outcomes

Complete interviews (Total)	111
Incomplete interviews	0
Eligible in process	0
Refusals	4
Out of target	95
Impossible to contact	106
Ineligible - coop.	4
Refusal to the Screener	85
Total	405

Appendix C

Madagascar Strict Universe Estimates

Region	Employees	Sector					Grand Total
		15	18	Other Manufacturing	52	Other Services	
Antananarivo	0-4	18	14	116	252	227	627
	5-19	21	19	127	186	173	525
	20-99	14	20	81	46	87	248
	100+	6	34	30	9	29	108
Antananarivo Total		59	86	354	494	515	1,508
Mahajanga	0-4			12	31	17	61
	5-19	2		12	16	13	44
	20-99	3		5	2	8	18
	100+			1		4	5
Mahajanga Total		5		30	50	42	128
Toamasina	0-4	4		13	53	28	99
	5-19	3		16	31	32	81
	20-99	2		11	10	13	36
	100+			4	2	5	11
Toamasina Total		8		44	96	78	226
Antsiranana	0-4			8	30	14	51
	5-19			8	14	14	36
	20-99	1		3	1	3	9
	100+					2	2
Antsiranana Total		1		19	45	33	98
Grand Total		74	86	446	685	668	1,960

Madagascar Weak Universe Estimates

Region	Employees	Sector					Grand Total
		15	18	Other Manufacturing	52	Other Services	
Antananarivo	0-4	52	44	247	643	761	1,747
	5-19	46	39	207	366	448	1,107
	20-99	24	49	126	86	213	498
	100+	17	103	60	22	85	286
Antananarivo Total		139	236	640	1,118	1,506	3,639
Mahajanga	0-4			22	68	48	139
	5-19	4		16	28	29	77
	20-99	3		6	3	17	29
	100+			2		7	10
Mahajanga Total		7		47	99	102	255
Toamasina	0-4	10		24	119	84	236
	5-19	5		23	53	72	152
	20-99	3		15	16	28	63
	100+			6	5	13	24
Toamasina Total		18		68	193	197	475
Antsiranana	0-4			13	64	39	117
	5-19			11	24	30	64
	20-99	3		4	2	7	16
	100+					5	5
Antsiranana Total		3		29	90	81	202
Grand Total		167	236	783	1,499	1,886	4,571

Madagascar Median Universe Estimates

Region	Employees	Sector					Grand Total
		15	18	Other Manufacturing	52	Other Services	
Antananarivo	0-4	32	25	169	475	441	1,143
	5-19	34	26	167	317	304	849
	20-99	19	35	110	81	157	403
	100+	13	74	51	21	61	220
Antananarivo Total		98	161	498	894	963	2,615
Mahajanga	0-4			18	59	33	110
	5-19	4		16	28	23	70
	20-99	3		6	3	15	27
	100+			2		6	8
Mahajanga Total		7		42	91	77	217
Toamasina	0-4	6		16	87	48	157
	5-19	4		18	45	48	115
	20-99	3		13	15	21	51
	100+			5	4	9	19
Toamasina Total		12		53	151	126	343
Antsiranana	0-4			10	53	25	88
	5-19			10	23	22	55
	20-99	2		4	2	6	14
	100+					4	4
Antsiranana Total		2		24	77	57	161
Grand Total		120	161	617	1,214	1,224	3,335

Appendix D

Madagascar Strict Cell Weights

Region	Employees	Sector				
		15	18	Other Manufacturing	52	Other Services
Antananarivo	0-4	1	2	11	21	19
	5-19	1	1	5	5	9
	20-99	1	1	2	1	3
	100+	1	2	2	1	1
Mahajanga	0-4			3	6	6
	5-19	1		3	4	3
	20-99	1		1	1	2
	100+			1		1
Toamasina	0-4	1		2	5	6
	5-19	1		2	3	5
	20-99	2		2	1	2
	100+			2	2	1
Antsiranana	0-4			2	4	3
	5-19			8	4	7
	20-99	1		3	1	2
	100+					1

Madagascar Weak Cell Weights

Region	Employees	Sector				
		15	18	Other Manufacturing	52	Other Services
Antananarivo	0-4	4	7	22	54	63
	5-19	3	2	7	10	22
	20-99	2	3	3	2	7
	100+	3	5	4	3	3
Mahajanga	0-4			6	14	16
	5-19	2		4	7	7
	20-99	1		1	1	4
	100+			2		2
Toamasina	0-4	2		3	12	17
	5-19	2		2	5	10
	20-99	3		3	2	4
	100+			3	5	3
Antsiranana	0-4			3	8	8
	5-19			11	6	15
	20-99	3		4	2	3
	100+					3

Madagascar Median Cell Weights

Region	Employees	Sector				
		15	18	Other Manufacturing	52	Other Services
Antananarivo	0-4	2	4	15	40	37
	5-19	2	1	6	8	15
	20-99	1	2	3	2	5
	100+	3	3	4	3	2
Mahajanga	0-4			4	12	11
	5-19	2		4	7	6
	20-99	1		1	2	4
	100+			2		2
Toamasina	0-4	1		2	9	10
	5-19	2		2	5	7
	20-99	3		2	2	3
	100+			3	4	2
Antsiranana	0-4			2	7	5
	5-19			10	6	11
	20-99	2		4	2	3
	100+					2