

The Madagascar 2022 Micro Enterprise Survey Implementation Report

I. Introduction

This document provides information on the Micro Enterprise Surveys (MES) implemented in Madagascar between March and December 2022. The MES collect both objective data based on firms' day-to-day experiences, and perceptions of enterprises regarding the business environment in which they operate.

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

II. Sampling Structure

The MES use stratified random sampling, where the population of establishments is first separated into non-overlapping groups, called strata, and then respondents are selected through simple random sampling from each stratum. The detailed methodology is provided in the Sampling Note.¹ Stratified random sampling has several advantages over simple random sampling. In particular, it:

- produces unbiased estimates of the whole population or universe of inference, as well as at the levels of stratification
- ensures representativeness by including observations in all of those categories
- produces more precise estimates for a given sample size or budget allocation, and
- may reduce implementation costs by splitting the population into convenient subdivisions.

The MES typically use two levels of stratification: industry classification, and subnational region (used in combination). Starting in 2022, the MES bases the industry classification on ISIC Rev. 4 (with earlier surveys using ISIC Rev. 3.1).² For regional coverage within a country, the MES has national coverage.

II.1 Stratification Categories

The Madagascar 2022 MES uses the following stratification categories:

- **Industry: 6 categories:**
 - Within manufacturing: Food, Garments, and Other Manufacturing

¹ The Sampling Note is available at: https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Sampling_Note-Consolidated-2-16-22.pdf. For further methodological background see, Richard L. Scheaffer; Mendenhall, W.; Lyman, R., "Elementary Survey Sampling", Fifth Edition.

² The WBES universe includes: all manufacturing (ISIC 4.0 codes 10-32), services (ISIC 4.0 codes 33, 41-43, 45-47, 49-56, 58, 61, 62, 69-75, 79, 95). Details on sectoral coverage and the WBES universe of inference can be found in the *Enterprise Surveys Manual and Guide* (p. 4).

- Within services: Wholesale, Retail, and Other Services
- **Size: 1 category:** Micro (1 to 4 employees).
- **Region: 8 categories:** Analamanga; Anosy; Atsimo Andrefana; Atsinanana; Boeni; Diana; Sava; and Vakinankaratra

Industry strata for the manufacturing sector were selected by their contribution to total value added using UNIDO INSTAT 4, available as of summer 2021. The manufacturing sectors stratified separately represent 16% of total value added with the rest of sectors grouped into a residual, "Other Manufacturing", stratum. By number of establishments, the manufacturing sectors represent 19% of the whole ES universe.

The strata for the services sector were selected by contribution to total number of establishments, using the Economic Census from INSTAT. The sectors stratified separately represent 46% of the total number of establishments in the services sector, and 37% of the whole ES universe.

Region strata of the Madagascar 2022 WBES were selected based on administrative divisions.

II.2 Universe

The universe of inference includes all formal (i.e., registered) private sector businesses (with at least 1% private ownership) and with at least five employees. In terms of sectoral criteria, all manufacturing businesses (ISIC Rev 3.1. codes 15-37) are eligible; for services businesses, those corresponding to the ISIC Rev 3.1 codes 45, 50-52, 55, 60-64, and 72 are included in the Micro Enterprise Surveys. Cooperatives and collectives are excluded from the Micro Enterprise Surveys. All eligible establishments must be registered with the registration agency. In the case of Madagascar, the definition of registration with INSTAT or EDBM was used. The universe table is the total number of eligible establishments, and the table is partitioned by the stratification groups (industry classification, establishment size, and subnational region) in a country.

For the Madagascar 2022 MES, the universe table, shown in Table 1 below, was obtained from INSTAT.

The frame obtained from INSTAT is the complete count of all establishments located within the geographical boundaries of the country. The sub-set corresponding to the standard ES universe was used.

II.3 Sampling Frame

The MES requires the most complete and up-to-date Sampling Frame, the list of eligible establishments with information on industry classification, size, address, and other contact information that will be used to randomly select the sample.

The Sampling Frame for the Madagascar 2022 MES was constructed from the following sources (see Table 2). There was no Panel Sampling Frame. The Fresh Sampling Frame, i.e., list of establishments that are in the MES universe, (ISIC 3.1 codes 15-37, 45, 50-52, 55, 60-64, and 72), was obtained from INSTAT.

For every MES, necessary measures are taken to ensure the quality of the frame; however, the sample frames are not immune to 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, eligibility adjustments may be needed when computing the appropriate sampling weights for individual observations (if and only if the weights are computed the sampling frame that is also the universe). Table 4 reports response outcomes.

II.4 Sample Design

The MES sample design, i.e., target number of interviews in each combination of stratification categories (cells), is generated using the two primary criteria: 1) minimize the difference from the purely proportional sample within each cell; and 2) achieve a sufficient sample size by stratification category to allow for estimates of a given level of precision.³ Additional information on the criteria for determining the sample size by stratification category is given in the *Sampling Note*, and additional information on the sample design is given in the *Enterprise Surveys Manual and Guide*.⁴ The original survey design for the Madagascar 2022 WBES is given in Table 3.

III. Data Collection

The detailed information on the WBES methodology and data collection is provided in the *Enterprise Surveys Manual and Guide*. The interviews for the Madagascar 2022 MES were conducted between March and December 2022. The interviews were conducted in the following languages: English, Malagasy, and French. For monetary variables, the currency was Malagasy Ariary.

III.1 Questionnaire

The questionnaire implemented in the Madagascar 2022 WBES included additional questions covering management practices, worker skills, and technology adoption. These questions were selected in collaboration with the members of the WB local country team.

III.2 Contractor

The fieldwork for the Madagascar 2022 MES was implemented by Dahlberg Research. The selection for the implementing agency followed the standard World Bank procurement practices that are described in more detail in the *Enterprise Surveys Manual and Guide*.

III.3 Sampling and screening

Samples are drawn by the Enterprise Surveys team in batches, following the stratification and sample design. The contractor conducted a thorough screening process before scheduling the

³ Additional constraints are also considered in the design stage. These are generally practical and include, for example, having a sufficient number of available contacts in the sample.

⁴ The *Enterprise Surveys Manual and Guide* is available at:

<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Enterprise-Surveys-Manual-and-Guide.pdf>

MES interviews. Results of the screening are provided using the eligibility and status codes as listed in Table 4. In cases of unit non-response (either a refusal or an inability to obtain an interview after exhaustive attempts), the contractor proceeded with the contact that appeared next in the list drawn in the respective cell. The process of sampling and screening is described in more detail in the *Enterprise Surveys Manual and Guide*.

III.4 Survey response

In all surveys, including the MES, some respondents choose not to participate. The Enterprise Analysis team and the contractor take all necessary measures to boost participation, through various methods of recruitment. The proper management of the screening process and sample replacement ensures that the resulting sample remains random.

The main measure of survey participation is the yield, which is the ratio of the total number of achieved interviews to the total number of contacted establishments. There are two main elements that boost yields. First is the survey participation rate, measured as the share of establishments that participated among those that can be assumed to have been eligible. The second element is the quality of frame. If only a small fraction of the contacted establishments is actually eligible to participate in the survey, then the sampling frame is far from ideal. This quality is measured by the ratio of the total number of establishments that can be assumed to be eligible for the WBES with the total number of contacted establishments in the frame. In other words:

$$Yield = Survey\ response\ rate * Rate\ of\ quality\ of\ the\ frame$$

which can be rewritten as follows:

$$\frac{Total\ interviewed}{Total\ contacted} = \frac{Total\ interviewed}{Assumed\ eligible} * \frac{Assumed\ eligible}{Total\ contacted}$$

Table 5 provides these measures for the Madagascar 2022 MES and across its stratification levels.

III.5 Achieved Sample

Tables 6 provide count of the MES interviews collected for each stratification cell, i.e., broken down by industry, establishment size, and region.

III.6 Sampling Weights

Since the MES uses stratified random sampling, individual observations should be properly weighted when making inferences about the population, since unweighted estimates are biased unless sample sizes are proportional to the size of each stratum. For each MES, special care is given to correctly compute sampling weights. Whenever the Universe is used to draw the sample (i.e., Sampling Frame is the same as the Universe) it is imperative to accurately adjust the universe within each stratum to account for the presence of ineligible establishments (e.g., the firm discontinued businesses, or is deemed ineligible due to its business activity or having fewer than five employees). Proper treatment of panel establishments is also crucial. Details about how the WBES sampling weights are calculated are given in the *Sampling Note*.

Three versions of sampling weights are calculated, depending on the assumptions that determine eligibility of establishments to be counted towards the WBES universe. These assumptions are called weak, medium, and strong; and are defined as given in the table below. Prevalence of each of these outcomes for the Madagascar 2022 MES is given in Table 4. All indicators and analysis conducted by the Enterprise Surveys team use the sampling weights based on the median assumption. Tables 7-9 report estimated universe based on the respective assumption.

Assumption	Eligibility codes for inclusion in the WBES universe
Strict	1,2,3,4,16
Median	1,2,3,4,16,10,11,13
Weak	1,2,3,4,16,10,11,13,91,92,93,94,12

For descriptions of each eligibility code, see Table 4.

III.7 Item response rates

Item response rate must be differentiated from survey response rates. The latter refers to participation in the survey itself (see Section III.4) whereas the former refers to the absence of responses to specific survey questions.⁵ The WBES, as any survey, suffer from item non-response; and different strategies are used by the data collection team to address this. In particular:

- For sensitive questions, such as on corruption or tax evasion, enumerators were instructed to collect the refusal to respond (-8) as a separate response category from don't know (-9).
- Establishments with incomplete information were re-contacted to fill gaps.

Table 10 provides item response rates for several key variables, broken down across the stratification levels.

III.8 Database Structure

The MES datafiles are organized in ways that reflect the corresponding questionnaire. The variables that are standard across countries have the first letter in their name correspond to the questionnaire section where the variable belongs in the questionnaire, i.e., *a1* denotes section A. The variables that are country or region specific are generally named using a prefix in capital letters. These country-specific variables are generally not available beyond the specific country or region where they were implemented. For the Madagascar 2022 MES, these variables use prefixes MG. 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.

The MES datafiles contain two establishment identifiers, *idstd* and *id*. The former is a global unique identifier of each establishment, while the latter is unique identifier within each survey. The variable *idstd* can be used to match the WBES establishment one-to-one across

⁵ The WBES questionnaire is organized so that there is always some entry in the database when the question was posed. An empty entry means that the question was not asked to the corresponding respondent, typically, due to skip patterns, or lack of applicability of that question in general.

databases. The variables *wweak*, *wmedian*, and *wstrong* corresponding to sampling weights based on, respectively, weak, median, and strong assumptions about eligibility (see Section III.6). The variable *strata* corresponds to the stratum of each observation.

Additionally, the MES datafiles contain many standard variables. The variable *d1a2_v4* denotes the main activity of the establishment, as obtained during the WBES interview, coded in the four-digit ISIC Rev. 4. Users should note that this activity may differ from the industry classification given in the Sampling Frame, *a4a*. Users are generally advised to use industry categories based on the realized information in *d1a2_v4*. Additional sampling information is contained in variables *a2* (region) and *a6a* (size). The variable *panel* identifies panel establishments, there are no panel establishments in Madagascar 2022. The combination of *a4a*, *a2*, *a6a*, and *panel* forms stratum of each establishment, which is contained in variable *strata*.

The last complete fiscal year for each establishment is contained in variables *a20m* (last month of last complete fiscal year) and *a20y* (last complete fiscal year).

Note that when an entry in the MES database is empty, this means that the question was not asked to the corresponding respondent. This happens when the question is deemed inapplicable, due to skip patterns or other reasons. In cases when the question was posed, some entry is provided, including don't know, which is an example of item non-response.

IV. Useful Links

The users of the WBES data may find the following links useful:

- *Sampling Note* is available at:
https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Sampling_Note-Consolidated-2-16-22.pdf
- The *Enterprise Surveys Manual and Guide* is available at:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Enterprise-Surveys-Manual-and-Guide.pdf>
- The WBES global questionnaires are available at:
<https://www.enterprisesurveys.org/en/methodology>
- The projects that are currently being implemented by the Enterprise Surveys team are available at: <https://www.enterprisesurveys.org/en/methodology/current-projects>
- The list of all WBES databases and detailed information about each is available here:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/DataDetails.xls>
- The description of the WBES indicators is available at:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/Indicator-Description.pdf>

References

- Cochran, William G., Sampling Techniques, New York, New York: John Wiley & Sons, 1977.
- Deaton, Angus, The Analysis of Household Surveys, Baltimore, Maryland: Johns Hopkins University Press, 1998.
- Levy, Paul S. and Stanley Lemeshow, Sampling of Populations: Methods and Applications, New York, New York: John Wiley & Sons, 1999.
- Lohr, Sharon L. Sampling: Design and Techniques, Boston, Massachusetts: Brooks/Cole, 1999.
- Scheaffer, Richard L.; Mendenhall, W.; Lyman, R., Elementary Survey Sampling, Fifth Edition, 1996.

Fact Sheet

Source of Universe Table	INSTAT
Source of Sampling Frame	INSTAT
Stratification sectors	Manufacturing of: Food, Garments, Other Manufacturing; Wholesale, Retail, and Other Services;
Stratification sizes	Small (5 to 19 employees), Medium (20 to 99 employees), Large (100 or more), and Unknown
Stratification regions	Analamanga; Anosy; Atsimo Andrefana; Atsinanana; Boeni; Diana; Sava; and Vakinankaratra
Contractor	Dahlberg Research
Fieldwork dates	March – December 2022
Interview languages	English, Malagasy, and French
Survey software	Survey Solution
Currency for nominal variables	Malagasy Ariary
Reference fiscal year	2021 and 2022
Sample Size	Total: Fresh: Panel:
Survey response rates	Yield: Response rate: Frame quality:
Item response rates	d2: n2a: 11: all TFP variables:
Additional topics covered in the questionnaire	Management practices, workers skills, and technology adoption
Additional surveys available (if any)	Micro Enterprise Survey, covering nine cities

Tables

Table 1: Madagascar 2022 WBES Universe

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Grand Total
Analamanga	Micro (1-4)	403	912	1,204	20,151	4,989	31,517	59,175
Anosy	Micro (1-4)	18	8	54	914	203	802	1,999
Atsimo Andrefana	Micro (1-4)	63	22	76	2,087	296	787	3,330
Atsinanana	Micro (1-4)	59	28	220	3,227	2,227	3,917	9,679
Boeni	Micro (1-4)	33	62	183	3,452	411	3,508	7,647
Diana	Micro (1-4)	87	30	245	3,411	233	2,708	6,715
Sava	Micro (1-4)	12	3	145	1,077	67	1,168	2,472
Vakinankaratra	Micro (1-4)	205	66	443	3,662	921	3,075	8,372
		879	1130	2568	37982	9347	47483	99389

Source: *INSTAT*.

Table 2: Madagascar 2022 WBES Sample Frame (Fresh and Panel Combined)

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Grand Total
Analamanga	Micro (1-4)	380	686	1,098	13,866	4,446	21,227	41,703
Anosy	Micro (1-4)	18	8	54	798	203	680	1,761
Atsimo Andrefana	Micro (1-4)	63	18	69	1,588	276	693	2,707
Atsinanana	Micro (1-4)	58	27	212	1,949	2,232	2,885	7,363
Boeni	Micro (1-4)	31	47	173	1,913	396	2,630	5,190
Diana	Micro (1-4)	80	30	244	2,412	233	1,731	4,730
Sava	Micro (1-4)	13	2	151	350	69	326	911
Vakinankaratra	Micro (1-4)	205	66	443	3,651	919	3,051	8,335
		848	884	2444	26527	8774	33223	72700

Source: *INSTAT*

Table 3: Original Survey Design (Fresh and Panel Combined)

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Grand Total
Analamanga	Micro (1-4)	22	11	4	16	6	25	84
Anosy	Micro (1-4)	13	6	7	24	7	20	77
Atsimo Andrefana	Micro (1-4)	10	13	6	19	16	8	71
Atsinanana	Micro (1-4)	4	3	8	19	11	17	62
Boeni	Micro (1-4)	3	11	7	13	4	8	46
Diana	Micro (1-4)	12	16	9	21	5	13	76
Sava	Micro (1-4)	9	2	21	28	15	22	97
Vakinankaratra	Micro (1-4)	2	4	2	14	4	8	34
		75	66	64	154	68	120	547

Table 4: Response Outcomes

- Modified table, to include eligibility codes, counts and shares

		Totals	Rates relative to total contacted
Overall	Contacts available in frame	72,700	
	Issued	5,031	
	Contacted	2,964	
Screening phase	Eligibles	1,853	62.5%
	Assumed eligibles	1,918	64.7%
	Screener refusals	24	0.8%
	Ineligible + out of target	134	4.5%
	Unobtainables	908	30.6%
Interview phase	Interview refusals	134	4.5%
(only if eligible)	Complete interviews	612	20.6%

Table 5: Survey Yield Rates

Stratification	Yield	Survey response rate	Frame quality
Fresh	21.0%	31.9%	65.7%
Micro (1-4)	21.0%	31.9%	65.7%
Analamanga	14.5%	22.1%	65.3%
Anosy	21.5%	57.0%	37.7%
Atsimo Andrefana	25.8%	32.3%	80.0%
Atsinanana	31.8%	36.2%	87.9%
Boeni	31.8%	62.5%	50.9%
Diana	31.8%	55.6%	57.2%
Sava	14.8%	18.6%	79.3%
Vakinankaratra	25.7%	47.9%	53.7%
Food	30.6%	57.9%	52.8%
Garments	30.2%	41.7%	72.4%
Other Manufacturing	26.2%	46.8%	56.0%
Retail	19.8%	31.3%	63.3%
Wholesale	18.4%	30.5%	60.4%
Other Services	17.1%	22.8%	74.8%
Madagascar 2022	21.0%	31.9%	65.7%

Notes: the rates are calculated as follows. Yield is 21.0%. Survey response rate is 31.9%. Frame quality is measured by 65.7%. By definition of these measures, yield is the product of survey response rate and frame quality.

Table 6: Achieved Total Sample (Fresh and Panel Combined)

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Grand Total
Analamanga	Micro (1-4)	25	20	4	19	7	51	126
Anosy	Micro (1-4)	4	2	13	22	17	19	77
Atsimo Andrefana	Micro (1-4)	7	12	7	20	15	10	71
Atsinanana	Micro (1-4)	6	6	9	20	12	15	68
Boeni	Micro (1-4)	4	9	8	18	4	12	55
Diana	Micro (1-4)	11	12	11	27	9	20	90
Sava	Micro (1-4)	4	1	20	30	13	22	90
Vakinankaratra	Micro (1-4)	5	6	2	13	2	7	35
		66	68	74	169	79	156	612

Table 7: Weak Universe Estimates

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Garments and Other Mfg	Grand Total
Analamanga	Micro (1-4)	362	849	1099	19454	4798	30570	0	57132
Anosy	Micro (1-4)	17	8	52	915	205	797	0	1993
Atsimo Andrefana	Micro (1-4)	42	16	54	1630	222	592	0	2556
Atsinanana	Micro (1-4)	52	26	199	3103	2111	3741	0	9231
Boeni	Micro (1-4)	26	50	147	2831	348	2908	0	6309
Diana	Micro (1-4)	78	28	225	3268	225	2609	0	6433
Sava	Micro (1-4)	11	2	4	1056	66	1175	147	2460
Vakinakaratra	Micro (1-4)	184	62	412	3558	900	2941	0	8057
		771	1040	2192	35815	8874	45333	147	94172

Table 8: Median Universe Estimates

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Garments and Other Mfg	Grand Total
Analamanga	Micro (1-4)	207	753	674	12977	3340	23120	0	41072
Anosy	Micro (1-4)	4	3	13	293	53	316	0	682
Atsimo Andrefana	Micro (1-4)	33	19	46	1268	216	520	0	2102
Atsinanana	Micro (1-4)	37	29	155	2514	1862	3235	0	7832
Boeni	Micro (1-4)	10	32	57	1662	148	1515	0	3424
Diana	Micro (1-4)	37	19	104	1714	119	1613	0	3607
Sava	Micro (1-4)	6	2	4	947	38	1146	82	2225
Vakinakaratra	Micro (1-4)	85	48	208	1539	528	1529	0	3936
		419	906	1261	22915	6302	32993	82	64879

Table 9: Strict Universe Estimates

		Food	Garments	Other Manufacturing	Retail	Wholesale	Other Services	Garments and Other Mfg	Grand Total
Analamanga	Micro (1-4)	185	716	619	12047	3211	22400	0	39178
Anosy	Micro (1-4)	4	3	13	280	55	321	0	675
Atsimo Andrefana	Micro (1-4)	30	20	45	1195	223	533	0	2046
Atsinanana	Micro (1-4)	34	30	149	2418	1906	3271	0	7809
Boeni	Micro (1-4)	7	28	45	1510	121	1340	0	3051
Diana	Micro (1-4)	32	18	89	1560	108	1550	0	3356
Sava	Micro (1-4)	6	2	4	946	39	1144	87	2227
Vakinakaratra	Micro (1-4)	65	41	166	1150	451	1304	0	3176
		364	856	1131	21106	6113	31861	87	61519

Table 10: Item Response Rates

Stratification	d2	l1	n2a	all TFP variables
Fresh	90.7%	99.8%	90.8%	0.0%
Micro (1-4)	90.7%	99.8%	90.8%	0.0%
Analamanga	82.4%	100.0%	88.0%	0.0%
Anosy	100.0%	100.0%	98.6%	0.0%
Atsimo Andrefana	100.0%	100.0%	97.0%	0.0%
Atsinanana	88.2%	100.0%	88.6%	0.0%
Boeni	96.4%	100.0%	90.0%	0.0%
Diana	88.9%	100.0%	88.7%	0.0%
Sava	98.9%	98.9%	90.9%	0.0%
Vakinankaratra	60.0%	100.0%	57.1%	0.0%
Food	71.4%	100.0%	78.9%	0.0%
Garments	86.8%	100.0%	80.0%	0.0%
Other Manufacturing	94.2%	100.0%	86.8%	0.0%
Madagascar 2022	90.7%	99.8%	90.8%	0.0%

Notes: the rates are calculated as follows. The number of responses with codes "-9" (don't know) or "-8" (refusal) in the respective variable (or combination of variables in the case of TFP) is divided by total number of respondents of the survey.