

The Sao Tome and Principe 2025 World Bank Enterprise Survey Implementation Report

I. Introduction

This document provides information on the World Bank Enterprise Surveys (WBES) implemented in Sao Tome and Principe between July and October 2025. The WBES collect both objective data based on firms' day-to-day experiences, and perceptions of enterprises regarding the business environment in which they operate. The WBES has global coverage. The WBES are also used to build a panel of enterprise data that will make it possible to track changes in the business environment over time.¹

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 WBES 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.² 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 than simple random sampling for a given sample size or budget allocation, and
- may reduce implementation costs by splitting the population into convenient subdivisions.

The WBES typically use three levels of stratification: industry classification, establishment size, and subnational region (used in combination). Starting in 2025, the WBES bases the industry classification on ISIC Rev. 4 (with earlier surveys using ISIC Rev. 3.1).³ For regional coverage within a country, the WBES has national coverage.

¹ A “panel interview” refers to an interview with a business that was also interviewed in the previous WBES.

² The Sampling Note is available at:

https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Sampling_Note-Consolidated-2-16-22.pdf.

³ The WBES universe includes: all manufacturing (ISIC 4.0 codes 10-33), services (ISIC 4.0 codes 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).

II.1 Universe

The universe of inference includes all formal (i.e., registered) private sector establishments⁴ (with at least 1% private ownership) and with at least five employees. In terms of sectoral criteria, all manufacturing businesses (ISIC Rev 4. codes 10-33) are eligible; for services businesses, those corresponding to the ISIC Rev 4 codes 41-43, 45-47, 49-53, 55-56, 58, 61-62, 69-75, 79, and 95 are included in the Enterprise Surveys. Cooperatives and collectives are excluded from the Enterprise Surveys. All eligible establishments must be registered with the registration agency. In the case of Sao Tome and Principe, the definition of registration by the Tax Directorate (*Direção dos Impostos - DI*) 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 Sao Tome and Principe 2025 WBES, the universe table, shown in Table 1 below, was obtained from 2020 Business Census (RE2020) provided the National Statistical Office (*Instituto Nacional de Estatística - INE*). The completeness of the Universe Table (and the Sampling Frame) was checked by cross-validating business register of establishments with data from the Tax Directorate (*Direção dos Impostos - DI*).

II.2 Sampling Frame

The WBES 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. In countries where a previous round of the WBES exists, the Sampling Frame also includes the Panel Sampling Frame, which provides information about all the establishments that participated in the previous round of the WBES in the country.

The Sampling Frame for the Sao Tome and Principe 2025 WBES was constructed from the following sources (see counts of establishments in the frame in Table 2). The Panel Sampling Frame was constructed using information on all the establishments that participated in the Sao Tome and Principe 2014 WBES. The Fresh Sampling Frame, i.e., list of establishments that are in the WBES universe and that have not participated in the Sao Tome and Principe 2014 WBES, for the manufacturing sector (ISIC 4.0 codes 10-32), was obtained mainly from the 2020 Business Census. Additional efforts were made to cross-validate business register and contact information of establishments with data from the business registry of the Tax Directorate (*Direção dos Impostos - DI*).

For every WBES, 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

⁴ The unit of analysis of every WBES is the establishment defined as a business entity associated with a physical location with its own set of financial statements, including Balance Sheet and Income Statement. However, in many economies official statistics are provided at the firm level, where a firm may be composed of several establishments, the so called multi-establishment firms.

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 universe equals the sampling frame (that is the weights project to universe totals by cell, coming from the sampling frame). Table 4 reports response outcomes.

II.3 Stratification Categories

The Sao Tome and Principe 2025 WBES uses the following stratification categories:

- **Industry: 2 categories:** Manufacturing and Services.
- **Size: 4 categories:** Small (5 to 19 employees), Medium (20 to 99 employees), Large (100 or more employees), Unknown
- **Region: 1 category:** Sao Tome and Principe

Industry strata for the manufacturing and services sectors were selected because the sample target is 150 observations.

II.4 Sample Design

The WBES 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. 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*.⁵ The original survey design for the Sao Tome and Principe 2025 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 Sao Tome and Principe 2025 WBES were conducted between July and October 2025. The interviews were conducted in Portuguese. For monetary variables, the currency was Dobra.

III.1 Questionnaire

The standard WBES questionnaire covers several topics regarding the business environment and business performance. These topics include general establishment characteristics, infrastructure, sales and supplies, trade, management practices, competition, innovation, capacity, land and permits, finance, business-government relations, exposure to bribery, labor, and performance. Information about the general structure of the questionnaire is available in the *Enterprise Surveys Manual and Guide*.

⁵ The *Enterprise Surveys Manual and Guide* is available at: <https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Enterprise-Surveys-Manual-and-Guide.pdf>

III.2 Contractor

The fieldwork for the Sao Tome and Principe 2025 WBES was implemented by *BELAB – Bissau Economics Lab* in collaboration with *TESE – Associação para o Desenvolvimento*. 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 ES 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 WBES, 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 are eligible for the WBES to 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 Sao Tome and Principe 2025 WBES and across its stratification levels.

III.5 Achieved Sample

Tables 6 and 7 provide counts of the WBES interviews collected for each stratification cell, i.e., broken down by industry, establishment size, and region. Table 6 reports full sample.

III.6 Sampling Weights

Since the WBES 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 WBES, 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 establishment 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 Sao Tome and Principe 2025 WBES 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 8-10 report estimated universe based on the respective assumption.

Assumption	Eligibility codes for inclusion in the WBES universe
Strict	1, 2, 3, 4
Median	1, 2, 3, 4, 10, 11, 13, 133
Weak	1, 2, 3, 4, 10, 11, 13, 133, 91, 92, 93, 94, 12

For descriptions of each eligibility code, see Table 4 of the sampling note.

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

III.8 Database Structure

The WBES 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*. 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 WBES 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 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.⁶

Additionally, the WBES 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, i.e., those that participated in the Sao Tome and Principe 2014 WBES. 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 WBES 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..

⁶ Note that the sampling weights may vary within *strata* for panel establishments due to the procedure used for projecting the universe from the previous round of the survey.

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/project-schedule>
- The list of all WBES databases and detailed information about each is available here:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/DataDetails.xls>
- The description of the WBES indicators is available at:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Indicator-Description.pdf>

Fact Sheet

Source of Universe Table	2020 Business Census (RE2020) provided the National Statistical Office (Instituto Nacional de Estatística - INE) & Tax Directorate Business registry (Direção dos Impostos - DI).		
Source of Sampling Frame	2020 Business Census (RE2020) provided the National Statistical Office (Instituto Nacional de Estatística - INE) & Tax Directorate Business registry (Direção dos Impostos - DI).		
Levels of Universe Table and Sampling Frame	Universe Table: Firm-level Sampling Frame: Firm-level		
Registration agency	Direção dos Impostos - DI		
Stratification sectors	Manufacturing and Services		
Stratification sizes	Small (5 to 19 employees), Medium (20 to 99 employees), Large (100-199), and Unknown		
Stratification regions	Sao Tome and Principe		
Contractor	BELAB – Bissau Economics Lab & TESE – Associação para o Desenvolvimento		
Fieldwork dates	July 2025 – October 2025		
Interview languages	Portuguese		
Survey software	Survey Solutions		
Currency for nominal variables	Dobra		
Reference fiscal year	2024		
VAT Applicable	YES		
Additional definitions	“Consumption Tax” was referred to as IVA.		
Sample Size	Total: 164	Fresh: 164	Panel: 0
Survey response rates	Yield: 20.4% Response rate: 69.9% Frame quality: 29.1%		
Item response rates	d2: 86.6% n2a: 89.6% 11: 94.4% all TFP vars : 66.7%		
Additional topics covered in the questionnaire	N/A		
Additional surveys available (if any)	N/A		

Tables

Table 1: Sao Tome and Principe 2025 WBES Universe

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	42	250	1,493
Sao Tome and Principe	Medium (20-99)	11	33	
Sao Tome and Principe	Large (100 or more)	2	6	
Sao Tome and Principe	Unknown	60	1089	
		115	1,378	1,493

Table 2: Sao Tome and Principe 2025 WBES Sample Frame (Fresh and Panel Combined)

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	42	250	1493
Sao Tome and Principe	Medium (20-99)	11	33	
Sao Tome and Principe	Large (100 or more)	2	6	
Sao Tome and Principe	Unknown	60	1089	
		115	1,378	1,493

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

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	27	38	150
Sao Tome and Principe	Medium (20-99)	9	20	
Sao Tome and Principe	Large (100 or more)	2	2	
Sao Tome and Principe	Unknown	5	47	
		43	107	150

Table 4: Response Outcomes

		Totals	Rates relative to total contacted
Overall	Contacts available in frame	1,493	
	Issued	1,169	
Screening phase	Contacted	810	
	Eligibles	236	29.1%
	Screener refusals	0	0.0%
	Assumed eligibles	236	29.1%
	Ineligible + out of target	369	45.6%
Interview phase	Unobtainables	205	25.3%
	Interview refusals	63	7.8%
(only if eligible)	Complete interviews	165	20.4%

Table 5: Survey Yield Rates

Stratification		Yield	Survey response rate	Frame quality
Panel	Fresh	20.4%	69.9%	29.1%
	Panel	#N/A	#N/A	#N/A
Size	Small (5-19)	31.4%	69.5%	45.2%
	Medium (20-99)	56.8%	80.6%	70.5%
	Large (100-199)	#N/A	#N/A	#N/A
	Very large (200+)	#N/A	#N/A	#N/A
Region	Sao Tome and Principe	20.4%	69.9%	29.1%
Sector	Manufacturing	21.0%	72.4%	29.0%
	Services	20.3%	69.6%	29.2%
Overall	Sao Tome and Principe 2025	20.4%	69.9%	29.1%

Notes: the rates are calculated as defined in Section III.4.

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

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	17	107	164
Sao Tome and Principe	Medium (20-99)	2	34	
Sao Tome and Principe	Large (100 or more)	0	4	
		19	145	164

Table 7: Weak Universe Estimates

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	73	568	734
Sao Tome and Principe	Medium (20-99)	4	72	
Sao Tome and Principe	Large (100 or more)	0	17	
		76	658	734

Table 8: Median Universe Estimates

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	33	243	336
Sao Tome and Principe	Medium (20-99)	3	49	
Sao Tome and Principe	Large (100 or more)	0	9	
		35	300	336

Table 9: Strict Universe Estimates

		Manufacturing	Services	Grand Total
Sao Tome and Principe	Small (5-19)	33	243	336
Sao Tome and Principe	Medium (20-99)	3	49	
Sao Tome and Principe	Large (100 or more)	0	9	
		35	300	336

Table 10: Item Response Rates

Screener/Interview		d2	l1	n2a	all TFP variables
Panel	Fresh	86.6%	99.4%	89.6%	66.7%
	Panel	#N/A	#N/A	#N/A	#N/A
Size	Small (5-19)	89.5%	100.0%	91.1%	77.8%
	Medium (20-99)	77.8%	97.2%	83.3%	33.3%
	Large (100+)	75.0%	100.0%	100.0%	#N/A
Region	Sao Tome and Principe	86.6%	99.4%	89.6%	66.7%
Sector	Manufacturing	78.9%	100.0%	89.5%	70.0%
	Services	87.6%	99.3%	89.7%	50.0%
Overall					
1	Sao Tome and Principe 2025	86.6%	99.4%	89.6%	66.7%

Notes: the rates are calculated as follows. The number of responses with informative content in the respective variable (or combination of variables in the case of TFP) is divided by the same number plus the number of responses that are "don't know" (code -9) or refusal (code -8).