

The Tajikistan 2024 World Bank Enterprise Survey Implementation Report

I. Introduction

This document provides information on the World Bank Enterprise Surveys (WBES) implemented in Tajikistan between April 2024 and October 2024. 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 currently cover over 220,000 firms in 159 countries with a total of 366 surveys since 2006. 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 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 2022, 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. For further methodological background see, Richard L. Scheaffer; Mendenhall, W.; Lyman, R., "Elementary Survey Sampling", Fifth Edition, 1996.

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 Stratification Categories

The Tajikistan 2024 WBES uses the following stratification categories:

- **Industry: 3 categories:**
 - All manufacturing sectors aggregated as one category: Manufacturing
 - Within services: Retail, and Other services
- **Size: 3 categories:** Small (5 to 19 employees), Medium (20 to 99 employees), Large (100 or more employees)
- **Region: 3 categories:**
 - Dushanbe
 - Sughd Region
 - Rest of the Country: covering Districts under Republican Subordination, Khatlon Region, and Gorno-Badakhshan Autonomous Region

For a small economy with the sample size of 360, the only industry classification is manufacturing, retail and other services.

Region strata of the Tajikistan 2024 WBES were selected based on administrative divisions, with some divisions grouped together to achieve minimum required precision of estimates at the level of each stratification region. In particular, the following regions were grouped together under “Rest of the Country” stratum: Districts under Republican Subordination, Khatlon region, and Gorno-Badakhshan Autonomous Region.

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 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 Tax Committee under the Republic of Tajikistan. 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 Tajikistan 2024 WBES, the universe table, shown in Table 1 below, was obtained from the Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT).

II.3 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 Tajikistan 2024 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 Tajikistan 2019 WBES. The Fresh Sampling Frame, i.e., list of establishments that are in the WBES universe and that have not participated in the Tajikistan 2019 WBES was obtained from the Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT)

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 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 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.⁴ 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 Tajikistan 2024 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 Tajikistan 2024 WBES were conducted between April 2024 and October 2024. The interviews were conducted in the following languages: Tajik and Russian. For monetary variables, the currency was Tajikistani Somoni.

III.1 Questionnaire

The standard WBES questionnaire covers several topics regarding the business environment and business performance. These topics include general firm characteristics, infrastructure, sales and supplies, trade, management practices, competition, innovation, capacity,

⁴ 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.

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>

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*.

III.2 Contractor

The fieldwork for the Tajikistan 2024 WBES was implemented by M-Vector in collaboration with ACT LTD. 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 can be assumed to be eligible for the WBES with the total number of contacted establishments in the frame. In other words:

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

which can be rewritten as follows:

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

Table 5 provides these measures for the Tajikistan 2024 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, while Table 7 shows counts of only panel interviews.

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 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 Tajikistan 2024 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
Weak	1,2,3,4,10,11,13,91,92,93,94,12

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.⁶ 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.

⁶ 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.

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.⁷

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 Tajikistan 2019 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. In cases when the question was posed, some entry is provided, including don't know, which is an example of item non-response.

⁷ 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/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>

Fact Sheet

Source of Universe Table	Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT), as of November 1, 2023		
Source of Sampling Frame	Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT), as of November 1, 2023, and 2019 World Bank Enterprise Survey		
Levels of Universe Table and Sampling Frame	Universe Table: Establishment-level Sampling Frame: Establishment-level		
Registration agency	Tax Committee under the Republic of Tajikistan		
Stratification sectors	Manufacturing; Retail; and Other Services		
Stratification sizes	Small (5 to 19 employees), Medium (20 to 99 employees), and Large (100 or more employees)		
Stratification regions	Dushanbe; Sughd; and Rest of the Country (covering Districts under Republican Subordination, Khatlon Region, and Gorno-Badakhshan Autonomous Region)		
Contractor	M-Vector in collaboration with ACT LTD		
Fieldwork dates	April 2024 – October 2024		
Interview languages	Tajik, Russian		
Survey software	Survey Solutions		
Currency for nominal variables	Tajikistani Somoni		
Reference fiscal year	2023		
VAT Applicable	YES		
Additional definitions	“Consumption Tax” was referred to as “Андоz аз истеъмол.” “Occupancy Permit” was referred to as “Санади комиссияи давлатии қабул дар бораи ба истифода қарор додани иншооти сохташуда.” For both capital city (a3b) and main business city (a3c), Dushanbe was entered.		
Sample Size	Total: 364	Fresh: 278	Panel: 86
Survey response rates	Yield: 16% Response rate: 63% Frame quality: 26%		
Item response rates	d2: 90.1% n2a: 95.3% 11: 100% all TFP vars.: 77.4%		
Additional topics covered in the questionnaire	None		
Additional surveys available (if any)	None		

Tables

Table 1: Tajikistan 2024 WBES Universe

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	285	147	607	1,635
Dushanbe	Medium (20-99)	117	47	306	
Dushanbe	Large (100 or more)	33	13	80	
Sughd	Small (5-19)	388	126	452	1,344
Sughd	Medium (20-99)	176	17	131	
Sughd	Large (100 or more)	37	3	14	
Rest of the Country	Small (5-19)	316	148	571	1,457
Rest of the Country	Medium (20-99)	143	27	200	
Rest of the Country	Large (100 or more)	31	3	18	
		1,526	531	2,379	4,436

Source: Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT)

Table 2: Tajikistan 2024 WBES Sample Frame (Fresh and Panel Combined)

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	285	147	607	1,635
Dushanbe	Medium (20-99)	117	47	306	
Dushanbe	Large (100 or more)	33	13	80	
Sughd	Small (5-19)	388	126	452	1,344
Sughd	Medium (20-99)	176	17	131	
Sughd	Large (100 or more)	37	3	14	
Rest of the Country	Small (5-19)	316	148	571	1,457
Rest of the Country	Medium (20-99)	143	27	200	
Rest of the Country	Large (100 or more)	31	3	18	
		1,526	531	2,379	4,436

Source: Agency on Statistics under the President of the Republic of Tajikistan (TAJSTAT); 2019 WBES

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

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	4	21	7	120
Dushanbe	Medium (20-99)	14	15	21	
Dushanbe	Large (100 or more)	10	4	24	
Sughd	Small (5-19)	7	32	5	120
Sughd	Medium (20-99)	32	6	20	
Sughd	Large (100 or more)	12	1	5	
Rest of the Country	Small (5-19)	4	31	9	120
Rest of the Country	Medium (20-99)	27	9	23	
Rest of the Country	Large (100 or more)	10	1	6	
		120	120	120	360

Table 4: Response Outcomes

		Totals	Rates relative to total contacted
Overall	Contacts available in frame	4,436	
	Issued	2,289	
	Contacted	2,279	
Screening phase	Eligibles	364	16.0%
	Screener refusals	217	9.5%
	Assumed eligibles	581	25.5%
	Ineligible + out of target	440	19.3%
	Unobtainables	1,258	55.2%
Interview phase	Interview refusals	0	0.0%
(only if eligible)	Complete interviews	364	16.0%

Table 5: Survey Yield Rates

	Stratification	Yield	Survey response rate	Frame quality
Panel	Fresh	14.3%	59.7%	24.0%
	Panel	25.4%	74.8%	34.0%
Size	Small (5-19)	17.1%	64.1%	26.7%
	Medium (20-99)	15.1%	63.4%	23.8%
	Large (100+)	15.9%	54.4%	29.3%
Region	Dushanbe	14.4%	48.2%	29.9%
	Sughd	14.9%	73.0%	20.5%
	Rest of the Country	19.3%	74.5%	25.9%
Sector	Manufacturing	17.0%	65.5%	25.9%
	Retail	19.2%	61.1%	31.5%
	Other Services	13.3%	61.1%	21.8%
Overall	Tajikistan 2024	16.0%	62.7%	25.5%

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

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

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	9	25	10	122
Dushanbe	Medium (20-99)	21	6	38	
Dushanbe	Large (100 or more)	6	1	6	
Sughd	Small (5-19)	28	27	10	119
Sughd	Medium (20-99)	23	5	17	
Sughd	Large (100 or more)	7	0	2	
Rest of the Country	Small (5-19)	6	32	10	123
Rest of the Country	Medium (20-99)	27	4	29	
Rest of the Country	Large (100 or more)	8	2	5	
		135	102	127	364

Table 7: Achieved Panel Sample

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	5	0	4	25
Dushanbe	Medium (20-99)	2	2	5	
Dushanbe	Large (100 or more)	3	0	4	
Sughd	Small (5-19)	8	2	3	29
Sughd	Medium (20-99)	5	3	3	
Sughd	Large (100 or more)	3	0	2	
Rest of the Country	Small (5-19)	4	1	2	32
Rest of the Country	Medium (20-99)	9	2	8	
Rest of the Country	Large (100 or more)	4	1	1	
		43	11	32	86

Table 8: Weak Universe Estimates

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	286	143	436	1,262
Dushanbe	Medium (20-99)	98	38	184	
Dushanbe	Large (100 or more)	25	9	43	
Sughd	Small (5-19)	435	137	362	1,233
Sughd	Medium (20-99)	165	0	0	
Sughd	Large (100 or more)	31	0	0	
Sughd	Medium and Large (20+)	0	16	87	
Rest of the Country	Small (5-19)	303	138	392	1,117
Rest of the Country	Medium (20-99)	115	21	115	
Rest of the Country	Large (100 or more)	22	2	9	
		1480	505	1628	3612

Table 9: Median Universe Estimates

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	89	67	164	476
Dushanbe	Medium (20-99)	32	19	72	
Dushanbe	Large (100 or more)	9	5	19	
Sughd	Small (5-19)	78	37	78	256
Sughd	Medium (20-99)	31	0	0	
Sughd	Large (100 or more)	8	0	0	
Sughd	Medium and Large (20+)	0	5	20	
Rest of the Country	Small (5-19)	76	52	118	340
Rest of the Country	Medium (20-99)	32	9	37	
Rest of the Country	Large (100 or more)	9	2	5	
		364	196	512	1072

Table 10: Strict Universe Estimates

		Manufacturing	Retail	Other Services	Grand Total
Dushanbe	Small (5-19)	44	32	75	233
Dushanbe	Medium (20-99)	21	8	38	
Dushanbe	Large (100 or more)	6	2	8	
Sughd	Small (5-19)	60	28	56	199
Sughd	Medium (20-99)	23	0	0	
Sughd	Large (100 or more)	7	0	0	
Sughd	Medium and Large (20+)	0	5	19	
Rest of the Country	Small (5-19)	62	41	90	272
Rest of the Country	Medium (20-99)	29	6	29	
Rest of the Country	Large (100 or more)	8	2	5	
		260	124	320	704

Table 11: Item Response Rates

	Screeners/Interview	d2	l1	n2a	all TFP variables
Panel	Fresh	89.2%	100.0%	94.6%	72.2%
	Panel	93.0%	100.0%	97.7%	88.4%
Size	Small (5-19)	92.9%	100.0%	97.6%	72.9%
	Medium (20-99)	87.3%	100.0%	93.3%	78.3%
	Large (100+)	89.1%	100.0%	93.5%	84.0%
Region	Dushanbe	85.4%	100.0%	89.4%	68.6%
	Sughd	88.2%	100.0%	98.3%	78.6%
	Rest of the Country	96.7%	100.0%	98.4%	83.3%
Sector	Manufacturing	90.2%	100.0%	96.2%	77.4%
	Retail	86.7%	100.0%	95.2%	N/A
	Other Services	91.9%	100.0%	94.6%	N/A
Overall	Tajikistan 2024	90.1%	100.0%	95.3%	77.4%

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).