

The India 2025 World Bank Enterprise Survey Implementation Report

I. Introduction

This document provides information on the World Bank Enterprise Surveys (WBES) implemented in India between April 2025 and April 2026. 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 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.

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 workers. 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. For India 2025 WBES, the definition of registration followed the respective sources of the universe table (see below).⁵ The universe table is the total number of eligible establishments, and it is partitioned by the stratification groups (industry classification, establishment size, and subnational region) in a country.

For the India 2025 WBES, the universe table was constructed using the following sources:

- Annual Survey of Industries (ASI), for reference year 2023–24, totaling 226,982 establishments.
- The counts received from the Ministry of Micro, Small & Medium Enterprises (M/o MSME), based on the Udyam portal, as of April 6, 2026. After excluding sectors outside the WBES universe, the MSME list contained 48,002,992 establishments. Because MSME size definitions differ from those used in WBES, additional processing steps were required. In particular, MSME list defines size based on the investment in plant machinery and equipment, along with the annual turnover:
 - Establishments with investment in plant machinery and equipment not more than INR 25 million and annual turnover not more than INR 100 million are categorized as Micro,
 - Establishments with investment in plant machinery and equipment not more than INR 250 million and annual turnover not more than INR 1 billion are categorized as Small, and
 - Establishments with not more than INR 1.25 billion and annual turnover not more than INR 5 billion are categorized as Medium.

Notably, the WBES team was informed that these sources likely overlap, but information about the extent of this overlap was not provided.

II.1.1 Processing of the MSME list

During fieldwork, 4,098 establishments sampled from the Udyam database (obtained as part of the sampling frame, see below) were screened. Region-specific shares of ineligible establishments were estimated and removed from the Micro-size category of the MSME-based

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.

⁵ See page 210, section 14 ("Registration Information") of "All India Report of Sixth Economic Census" available here: <https://msme.gov.in/sites/default/files/All%20India%20Report%20of%20Sixth%20Economic%20Census.pdf>

universe. More specifically, average share of ineligible establishments was estimated at around 7.4%. This adjustment reduced the MSME universe to 44,659,835 establishments.

To avoid double counting, establishments already included in the sampling frame under the “unknown” size category (i.e., from Udyam source) were removed from the MSME counts, and specifically from the “Micro” size category (these are added back later through the frame-based component of the universe estimation).

Among the 4,098 sampled Udyam establishments, 2,566 were successfully interviewed. Turnover information from these interviews was used to map MSME size categories to WBES size classes. Establishments with annual turnover of up to INR 100 million were classified as Micro, those above INR 100 million and up to INR 1 billion as Small, and those above INR 1 billion and up to INR 5 billion as Medium. For each MSME size class and stratification region, the distribution across standard WBES size categories (5–19, 20–99, 100–199, and 200+ workers) was calculated and used to reallocate MSME counts accordingly.

II.1.2 Final Construction of the Universe Table

To combine establishment counts across all sources:

- ASI counts for 2023–24 were used as received,
- MSME counts as of April 6, 2026 were used after the adjustments described above,
- Sampling frame counts were incorporated directly, including Udyam establishments classified under the “unknown” size category.

For each cell, the maximum value across these three sources was taken, reflecting potential under-coverage of larger firms in the MSME data and the absence of data from the Ministry of Corporate Affairs (MCA, which is included in the Sampling Frame, see below) from the universe construction. The resulting universe table is shown on Table 1. While the Universe Table used in the India 2025 WBES has substantially larger number of entities than estimated based on the 2022 WBES (approximately 45 million compared with 2.5 million)—possibly reflecting the Udyam universe—the WBES team took as many statistically robust steps as possible, described above, to limit potential overestimation.

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 India 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 India 2022 WBES.

The Fresh Sampling Frame, i.e., list of establishments that are in the WBES universe and that have not participated in the India 2022 WBES, was constructed using multiple administrative

and survey-based sources, following a process of coverage assessment, cleaning, and de-duplication. In particular, the following sources were used:

- Annual Survey of Industries (ASI) – for the reference year 2022–2023. The original file contained 255,212 establishments. After excluding records from sectors outside the WBES universe, 252,061 establishments were retained following the first stage of cleaning.
- Ministry of Corporate Affairs (MCA) – the list of firms operating as of April 2025 was received from the Ministry. The original dataset contained 1,872,575 entries. Importantly, this list is firm-level, while the other sources are establishment-level. After excluding firms operating in sectors outside the WBES universe, 1,349,254 entries were retained following the initial cleaning stage.
- Annual Survey of Unincorporated Sector Establishments (ASUSE) – for reference year 2023-2024, list of registered unincorporated establishments was obtained. The raw file contained 20,287 entries. After removing exact duplicates, 20,181 establishments were retained.
- Udyam Registration Database - To complement the sample frame, establishment-level administrative data were extracted from the Government of India’s Udyam Registration portal. A custom script was developed in Python using Selenium and BeautifulSoup to retrieve publicly available registration records. The script iterated over a comprehensive set of five-digit National Industrial Classification (NIC-5) codes covering manufacturing, trade, and services activities within the scope of the World Bank Enterprise Surveys and queried all Indian states and union territories. For each NIC category and location, the script parsed the search results table and captured enterprise-level fields (including enterprise name, address, state, district, pin code, social category, and gender), along with the corresponding NIC industry definition. The data extraction was conducted between March and July 2025, and all records were stored in a structured CSV format for subsequent cleaning, validation, and analysis. This process resulted in extracting a total of 21,111,026 entries.

To assess eligibility for inclusion in the WBES universe, a random sample of 1,000 Udyam records was pre-screened. To minimize costs, screening was concentrated in five large urban areas—Delhi, Mumbai, Hyderabad, Kolkata, and Kamrup—corresponding to five large WBES fieldwork regions, with 200 random observations allocated to each. Based on pre-screening, the estimated shares of eligible establishments were: 44.5% in Delhi, 71% in Mumbai, 57% in Hyderabad, 59% in Kolkata, and 22% in Kamrup. These shares were applied to the corresponding fieldwork regions in the full Udyam dataset, resulting in 9,104,593 retained establishments. After excluding 464 panel matches, the final contribution from the Udyam database was 9,104,129 establishments. Since size information was not available in the publicly accessible portal, all of these entries enter in the size category “unknown”.

After de-duplication across the sources listed above, the final total number of entries in the Sampling Frame is 10,732,049. Table 2 shows counts across the stratification cells.

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 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 India 2025 WBES uses the following stratification categories:

- **Industry: 18 categories:**
 - Within manufacturing: Food, Textiles, Garments, Chemicals & Chemical Products, Pharmaceuticals & Medical Products, Rubber & Plastics Products, Non-Metallic Mineral Products, Basic Metals, Machinery & Equipment, Motor Vehicles, and Other Manufacturing.
 - Within services: Construction, Wholesale, Retail, Hotels, Transport, IT & IT Services, Other Services
- **Size: 5 categories:** Small (5 to 19 workers), Medium (20 to 99 workers), Large (100 to 199 workers), Very Large (200 or more workers), and Size-unknown.
- **Region: 24 categories:** Andhra Pradesh, A&N Islands; Arunachal Pradesh, Nagaland, Manipur, Tripura, Meghalaya, Mizoram, Sikkim; Assam; Bihar; Chhattisgarh; Delhi; Goa; Gujarat, Dadra & Nagar Haveli & Daman & Diu; Haryana; Himachal Pradesh; Jammu & Kashmir (union territory), Ladakh; Jharkhand; Karnataka; Kerala; Madhya Pradesh; Maharashtra; Odisha; Punjab, Chandigarh; Rajasthan; Tamil Nadu, Puducherry; Telangana; Uttar Pradesh; Uttarakhand; and West Bengal.

Industry strata for the manufacturing sector were selected by their contribution to total value added, to employment and to total number of establishments, using UNIDO INSTAT 4, available as of March 2025. The manufacturing sectors stratified separately represent 75%, 72% and 73% of total value added, employment, and number of establishments, respectively, with the rest of sectors grouped into a residual, "Other Manufacturing", stratum. Based on the universe table, by number of establishments, the manufacturing sectors stratified separately represent 21% of the whole ES universe.

The strata for the services sector were selected by estimated contribution to total number of establishments, using the India 2022 WBES sampling weights. The sectors stratified separately were estimated to represent 72% of the total number of establishments in the services sector, and 44% of the whole ES universe. Based on the universe table (which was not available at the time when the stratification sectors were selected), these shares are 69% and 46%, respectively.

Region strata of the India 2025 WBES were selected based on administrative divisions, with some territories grouped together to achieve minimum required precision of estimates at the

level of each stratification region. In particular, relatively small north-eastern states were grouped together (Arunachal Pradesh, Nagaland, Manipur, Tripura, Meghalaya, Mizoram, and Sikkim), and while the rest of the regions follow existing administrative divisions within India, several territories were combined to ensure full national coverage (e.g., Andaman and Nicobar Islands were combined with Andhra Pradesh). Such combinations are clear in the names that are used for the stratification regions.

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*.⁶ The original survey design for the India 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 fieldwork for the India 2025 WBES was conducted between April 2025 and April 2026. The interviews were offered in the following languages: Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Tamil, Telugu and English (but mostly Hindi and English). For monetary variables, the currency was Indian rupee (INR).

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*.

The India 2025 WBES contains two survey questions that did not appear in other surveys of the same rollout. In particular, question *IND1* on firms' agreements with their suppliers, and question *SAR131* on provision of childcare facilities was added.

III.2 Contractor

The fieldwork for the India 2025 WBES was implemented by NielsenIQ (India) Private Limited. The selection for the implementing agency followed the standard World Bank

⁶ The *Enterprise Surveys Manual and Guide* is available at:
<https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Enterprise-Surveys-Manual-and-Guide.pdf>

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:

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

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 India 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, 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 establishment discontinued businesses, or is deemed ineligible due to its business activity or having fewer than five workers). 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 India 2025 WBES is given in Table 4. For the India 2025 WBES, since the sources of the Universe Table and Sampling Frame are different, base weights were calculated, without applying these adjustments. The estimated weights still vary slightly across the assumptions, driven by the Sampling Weights of the panel establishments (i.e., those that also participated in the India 2022 WBES). 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. The estimated total size across the respective assumptions does not add up exactly to the total universe due to rounding across many stratification cells.

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. Additional questions were fielded in the India 2025 WBES, namely question *INDI* on firms’ agreements with their suppliers, and question *SAR131* on provision of childcare facilities.

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 India 2022 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.

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/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	Annual Survey of Industries (ASI) – reference year 2023–2024, and counts from the Ministry of Micro, Small & Medium Enterprises (MSME), based on the Udyam portal, as of April 6, 2026.
Source of Sampling Frame	Annual Survey of Industries (ASI) – reference year 2022–2023; the list from the Ministry of Corporate Affairs (MCA) as of April 2025; Annual Survey of Unincorporated Sector Establishments (ASUSE) –reference year 2023-2024; and Udyam Registration Database – scraped from the public portal between March and July 2025.
Levels of Universe Table and Sampling Frame	Universe Table: mix of establishment-level and firm-level Sampling Frame: mix of establishment-level and firm-level
Registration agency	Following definitions of the sources of the Universe Table and Sampling Frame
Stratification sectors	Manufacturing of: Food, Textiles, Garments, Chemicals & Chemical Products, Pharmaceuticals & Medical Products, Rubber & Plastics Products, Non-Metallic Mineral Products, Basic Metals, Machinery & Equipment, Motor Vehicles, and Other Manufacturing; Construction, Wholesale, Retail, Hotels, Transport, IT & IT Services, Other Services.
Stratification sizes	Small (5 to 19 workers), Medium (20 to 99 workers), Large (100-199), Very Large (200 or more workers), and Size-unknown
Stratification regions	Andhra Pradesh, A&N Islands; Arunachal Pradesh, Nagaland, Manipur, Tripura, Meghalaya, Mizoram, Sikkim; Assam; Bihar; Chhattisgarh; Delhi; Goa; Gujarat, Dadra & Nagar Haveli & Daman & Diu; Haryana; Himachal Pradesh; Jammu & Kashmir (union territory), Ladakh; Jharkhand; Karnataka; Kerala; Madhya Pradesh; Maharashtra; Odisha; Punjab, Chandigarh; Rajasthan; Tamil Nadu, Puducherry; Telangana; Uttar Pradesh; Uttarakhand; and West Bengal
Contractor	NielsenIQ (India) Private Limited
Fieldwork dates	April 2025 – April 2026
Interview languages	Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Tamil, Telugu and English
Survey software	Survey Solutions
Currency for nominal variables	Indian rupee (INR)
Reference fiscal year	2025
VAT Applicable	Yes

Additional definitions	“Consumption Tax” was referred to as GST. For the capital city (a3b), New Delhi was entered; for the main business city (a3c), Mumbai was entered.		
Sample Size	Total: 10,479	Fresh: 6,260	Panel: 4,219
Survey response rates	Yield: 90% Response rate: 96% Frame quality: 94%		
Item response rates	d2: 100% n2a: 100% 11: 100% all TFP vars. ⁸ : 100%		
Additional topics covered in the questionnaire	supplier agreements, childcare facilities		

Tables

Due to large size of the tables for the India 2025 WBES, they are provided in the excel file "India 2025_ES_Implementation Report Tables.xlsx", which is available along with this Implementation Report in the zipfile that contains all the documentation for this survey.

⁸ TFP stands for total factor productivity. TFP variables are d2 (annual sales), n2a (annual cost of labor), n2e (annual cost of intermediate goods) and n7e (replacement cost of capital). Response rate for TFP variables is calculated only for manufacturing establishments.