

The India 2022 World Bank Enterprise Survey Implementation Report

I. Introduction

This document provides information on the World Bank Enterprise Surveys (WBES) implemented in India between December 2021 and October 2022. 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 180,000 firms in 154 countries with a total of 306 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 10-32), 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).

II.1 Stratification Categories

The India 2022 WBES uses the following stratification categories:

- **Industry: 17 categories:**
 - Within manufacturing: Food, Textiles, Garments, Chemicals & Chemical Products, Non-Metallic Mineral Products, Basic Metals, Fabricated Metal Products, Machinery & Equipment, Motor Vehicles, Other manufacturing
 - Within services: Construction, Wholesale, Retail, Hotels, Restaurants, IT & IT Services, and Other Services
- **Size: 4 categories:** Small (5 to 19 employees), Medium (20 to 99 employees), Large (100-199), and Very Large (200 or more employees)
- **Region: 24 categories:** Andhra Pradesh; Arunachal Pradesh, Nagaland, Manipur, Tripura, and Meghalaya; Assam; Bihar; Chhattisgarh; Delhi; Goa; Gujarat; Haryana; Himachal Pradesh; Jammu & Kashmir (union territory); Jharkhand; Karnataka; Kerala; Madhya Pradesh; Maharashtra; Odisha; Punjab; Rajasthan; Tamil Nadu; 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 summer 2021. The manufacturing sectors stratified separately represent 60%, 68% and 70% of total value added, employment, and number of establishments, respectively, with the rest of sectors grouped into a residual, "Other Manufacturing", stratum. By number of establishments, the manufacturing sectors stratified separately represent 30% of the whole ES universe.

The strata for the services sector were selected by contribution to total number of establishments, using the 6th Economic Census. The sectors stratified separately represent 78% of the total number of establishments in the services sector, and 42% of the whole ES universe.

Region strata of the India 2022 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, and Meghalaya), with the rest of the 23 regions following existing administrative divisions within India.

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 registration agency. In the case of India, the definition of registration of the 6th Economic Census (EC) was used, where

registration can be from any of the following: Shops and Commercial Establishments Act; Companies Act, 1956; Factories Act, 1948; Central Excise/Sales Tax Act; Societies Registration Act; Co-operative Societies Act; Directorate of Industries; KVIC/KVIB/DC: Handloom/Handicrafts; Registered with other relevant agencies.⁴ 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 India 2022 WBES, the universe table, shown in Table 1 below, was obtained from the 6th EC, that was conducted between January 2013 and April 2014.⁵ The EC 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. The completeness of the Universe Table (and the Sampling Frame) was checked by cross-validating counts of establishments with the corresponding tables from the National Sample Survey Office (NSSO) and the annual report on the Micro, Small & Medium Enterprises (MSMEs, 2019-2020).

As discussed in Section III.6 in more detail, for calculation of sampling weights, the universe from the 6th EC was replaced with the universe constructed from the sampling frame.

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 India 2022 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 2014 WBES. The Fresh Sampling Frame, i.e., list of establishments that are in the WBES universe and that have not participated in the India 2014 WBES, for the manufacturing sector (ISIC 4.0 codes 10-32), was obtained mainly from the Annual Survey of Industries (ASI), 2018-2019, which includes establishments that are registered under the Factories Act of India. However, ASI excludes establishments employing fewer than 10 employees. To fill this gap, the implementing agency (Nielsen IQ) used the list of formally registered firms employing 5-9 workers from the 6th Economic Census. The contact information of this set of establishments (i.e., employing 5-9 workers, including those in services sectors, see on this below) was less readily available for Nielsen IQ. Additional efforts were made to collaborate with the Ministry of Statistics and Programme Implementation (MOSPI) to gain access to this information.

⁴ 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>

⁵ Attempts were made to obtain the universe table from the 7th Economic Census, which started in 2019. However, the data was not yet finalized or made available during the preparation phase of the India ES.

The Fresh Sampling Frame for the services sector (ISIC 4.0 codes 33, 41-43, 45-47, 49-56, 58, 61, 62, 69-75, 79, 95) is from the 6th EC (sub-sample corresponding to the ES services universe). Nielsen IQ verified this list with lists through publicly available sources for some business organizations to try to update the census data as much as possible, while avoiding introducing duplicates as much as possible.

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.⁶ 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*.⁷ The original survey design for the India 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 India 2022 WBES were conducted between December 2021 and October 2022. The interviews were conducted in the following languages: Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Tamil, Telugu and English (but mostly Hindi and English). For monetary variables, the currency was Indian rupee.

Apart from the challenges with participation and call-backs that are standard for all surveys, the India 2022 WBES faced challenges interviewing firms belonging to the Assam Branch Indian Tea Association (ABITA). Multiple attempts were made to obtain the permissions requested by the firms from the Secretary of ABITA and the relevant labor secretaries, but these attempts were not successful. These contacts were marked as refusal to screener in the progress report.

⁶ 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.

⁷ 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.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, 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 questionnaire implemented in the India 2022 WBES included additional questions covering contractual disputes, COVID-19, green economy, delayed payments, invoice discounting (TReDS or similar services), government support, attitudes towards taxes, training costs, and childcare support. These questions were selected in collaboration with the members of the WB local country team.

III.2 Contractor

The fieldwork for the India 2022 WBES was implemented by NielsenIQ (India) Private Limited. 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:

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

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 India 2022 WBES and across its stratification levels.

III.5 Achieved Sample

Tables 6 and 7 provide count 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 India 2022 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	1,2,3,4,16
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.

Note that for the India 2022 WBES, there were two potential ways to calculate sampling weights. First, the universe tables from the 6th EC can be used along with the frame constructed from other sources as described in section II.3, without applying the adjustments described above (since the universe and frame would be from separate sources). Second, the universe can be constructed from the frame, allowing the application of adjustments as described above. The counts of establishments from the two universe tables are within 0.5% of each other. The

comparison of the two methods showed that estimated universe for median assumption is within 9% of the unadjusted universe from the first method. The choice was made to use the second approach, to make direct use of the information about the status of establishments obtained through the screening process.

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.⁸ 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. 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 India 2022 WBES, these variables use prefixes IND, SAR, ASC, BMG, and COV. 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.

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

⁸ 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.

establishments, i.e., those that participated in the India 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. 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>

Fact Sheet

Source of Universe Table	6 th Economic Census (EC, conducted between January 2013 and April 2014)
Source of Sampling Frame	Manufacturing: Annual Survey of Industries (ASI), 2018-2019, supplemented from the 6 th EC Services: 6 th EC
Stratification sectors	Manufacturing of: Food, Textiles, Garments, Chemicals & Chemical Products, Non-Metallic Mineral Products, Basic Metals, Fabricated Metal Products, Machinery & Equipment, Motor Vehicles, Other manufacturing; Construction, Wholesale, Retail, Hotels, Restaurants, IT & IT Services, and Other Services
Stratification sizes	Small (5 to 19 employees), Medium (20 to 99 employees), Large (100-199), and Very Large (200 or more employees)
Stratification regions	Andhra Pradesh; Arunachal Pradesh, Nagaland, Manipur, Tripura, and Meghalaya; Assam; Bihar; Chhattisgarh; Delhi; Goa; Gujarat; Haryana; Himachal Pradesh; Jammu & Kashmir (union territory); Jharkhand; Karnataka; Kerala; Madhya Pradesh; Maharashtra; Odisha; Punjab; Rajasthan; Tamil Nadu; Telangana; Uttar Pradesh; Uttarakhand; and West Bengal
Contractor	NielsenIQ (India) Private Limited
Fieldwork dates	December 2021 – October 2022
Interview languages	Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Tamil, Telugu and English
Survey software	CSPPro
Currency for nominal variables	Indian rupee
Reference fiscal year	2021 (2,927 obs.) and 2022 (6,449 obs.)
Sample Size	Total: 9,376 Fresh: 5,310 Panel: 4,066
Survey response rates	Yield: 57% Response rate: 62% Frame quality: 92%
Item response rates	d2: 99.8% n2a: 99.4% 11: 99.2% all TFP vars.: 95.2%
Additional topics covered in the questionnaire	contractual disputes, COVID-19, green economy, delayed payments, invoice discounting (TReDS or similar services), government support, attitudes towards taxes, training costs, and childcare support
Additional surveys available (if any)	Informal Sector Enterprise Survey, Micro Enterprise Survey, each covering nine cities

Tables

Due to large size of the tables for the India 2022 WBES, they are provided in the excel file "ES Implementation Report Tables_India 2022.xlsx", which is available along with this Implementation Report in the zipfile that contains all the documentation for this survey.