

The United States 2024 World Bank Enterprise Survey Implementation Report

I. Introduction

This document provides information on the World Bank Enterprise Surveys (WBES) implemented in the United States between May 2024 and April 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 currently cover over 232,000 firms in 162 countries with a total of 385 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 United States 2024 WBES uses the following stratification categories:

- **Industry: 16 categories:**
 - Within manufacturing: Food; Garments; Pharmaceutical, Chemicals & Chemical Products; Fabricated Metal Products; Computers & Electronics; Machinery & Equipment; Motor Vehicles; and Other Manufacturing
 - Within services: Construction; Wholesale; Retail; Hotels; Restaurants; Professional Activities; and Other Services
- **Size: 4 categories:** Small (5 to 19 employees), Medium (20 to 99 employees), Large (100 to 249 employees), and Very large (250 or more)
- **Region: 14 categories:** New England, Middle Atlantic (excl. New York MSA), New York Metro Statistical Area, South Atlantic (excl. DC MSA), DC Metro Statistical Area, East South Central, East North Central (excl. Chicago MSA), Chicago Metro Statistical Area, West North Central, West South Central (excl. Dallas MSA), Dallas Metro Statistical Area, Mountain, Pacific (excl. Los Angeles MSA), Los Angeles Metro Statistical Area

Industry strata for the manufacturing sector were selected by their contribution to total value added, to employment and to total number of establishments, using the UNIDO INSTAT and the County Business Patterns (CBP) data, released in March 2023. According to the UNIDO data on value added, the seven selected categories in manufacturing account for over 60% of value added in manufacturing; the sectors other than garments manufacturing also represent the six largest in terms of value added. Garment manufacturing was selected for international comparability purposes. The strata for the services sector were selected by contribution to total number of establishments and employment (that is, within the WBES universe), using the CBP data. The sectors stratified separately represent 79% of establishments and 77% of employment within the WBES services sectors (and 70% of establishments and 63% of employment within the full WBES universe).

Region strata of the United States 2024 WBES were selected based on statistical regions used by the U.S. Census Bureau⁴; in the case of the urban areas of New York, Washington, DC, Chicago, Dallas, and Los Angeles, additional metropolitan statistical areas (MSAs)⁵ were separately stratified. The census regions where these MSAs were originally included, therefore, denote when any such MSA is excluded, e.g., “Middle Atlantic (excl. New York MSA)”.

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

⁴ Found here: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

⁵ Delineations found here: <https://www.census.gov/programs-surveys/metro-micro/geographies/geographic-reference-files.html>

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 the United States, registration was considered as being registered with the Business Registry and/or the Internal Revenue Service, as a business entity.

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 United States 2024 WBES, the universe table, shown in Table 1 in the accompanying Excel document, was obtained from the CBP data, available as of March 2023.

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 Fresh Sampling Frame, i.e., list of establishments that are in the WBES universe and that have not participated in the United States 2024 WBES, for all establishments was obtained from a private commercial source, Orbis, provided by Bureau Van Dijk. The full Orbis data used for the frame was available as of December 2023. Enrichment exercises, including the enrichment of records through Google was systematically performed.

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). However, in cases where the universe table and sampling frame are from different sources, as is the case with the United States 2024 WBES, such adjustments are not appropriate.

Table 4 in the accompanying Excel document 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

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.

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 United States 2024 WBES is given in Table 3 in the accompanying Excel document.

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 United States 2024 WBES were conducted between May 2024 and April 2025. The interviews were conducted primarily in English and (occasionally) Spanish. For monetary variables, the currency was US dollar (USD).

Aside from the typical difficulties of data collection, response rates were noticeably lower among garment manufacturers. For this reason, for public release, garment manufacturers are presented in “Other Manufacturing”.

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, 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 United States 2024 WBES was implemented by Kapa Research, in collaboration with NORC at the University of Chicago and the Global Data Collection Company (GDCC). 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

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>

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 in the accompanying Excel document provides these measures for the United States 2024 WBES and across its stratification levels.

III.5 Achieved Sample

Tables 6 and 7 in the accompanying Excel document 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 United States 2024 WBES is given in Table 4 in the accompanying Excel document. 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.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. 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.⁹

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.

⁹ 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.

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 United States 2024 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	CBP available as of March 2023
Source of Sampling Frame	Orbis available as of December 2023
Levels of Universe Table and Sampling Frame	Universe Table: Establishment-level Sampling Frame: Establishment-level
Registration agency	Business Register and/or IRS
Stratification sectors	Manufacturing of: Food; Garments; Chemicals & Chemical Products; Basic Metals; Electrical & Computer Products; Electronics; Machinery & Equipment; Motor Vehicles; Other Manufacturing; Retail; Construction; Professional Activities; Other Services
Stratification sizes	Small (5 to 19 employees), Medium (20 to 99 employees), Large (100-249), and Very Large (250 or more employees)
Stratification regions	New England; Middle Atlantic (excl. New York MSA); New York Metro Statistical Area; South Atlantic (excl. DC MSA); DC Metro Statistical Area; East South Central; East North Central (excl. Chicago MSA); Chicago Metro Statistical Area; West North Central; West South Central (excl. Dallas MSA); Dallas Metro Statistical Area; Mountain; Pacific (excl. Los Angeles MSA); Los Angeles Metro Statistical Area
Contractor	Kapa Research, NORC (at the University of Chicago), and GDCC
Fieldwork dates	May 2024 – April 2025
Interview languages	English, Spanish
Survey software	Survey Solutions
Currency for nominal variables	USD
Reference fiscal year	2023 and 2024
VAT Applicable	NO
Additional definitions	For the capital city (a3b), Washington, DC was entered; for the main business city (a3c), New York City was entered.
Sample Size	Total: 2,697
Survey response rates	Yield: 5.7% Response rate: 11.5% Frame quality: 50%

Item response rates	d2: 100% n2a: 99.6% 11: 100% all TFP vars.: 84.9%
Additional topics covered in the questionnaire	N/A
Additional surveys available (if any)	N/A

Tables can be found in the accompanying Excel document.