MSME Country Indicators 2014

DESCRIPTION NOTE

Erick Gonzales
Martin Hommes
Melina Laura Mirmulstein

IFC International Finance Corporation
WORLD BANK GROUP

DECEMBER 2014

Contents

Terms of Use and Disclaimer

Please read the "Description Note on the MSME Country Indicators 2014" along with the country specific comments in the Micro, Small, and Medium Enterprise Country Indicators (MSME-CI) MS Excel workbook before using the data. MSME-CI presents secondary data collected by various institutions (statistical institutes, ministries, international organizations, small business promotion agencies, research institutions and others) using different methods (survey, census and others). IFC is not responsible for the quality, accuracy, reliability or completeness of the data these sources provide. Data are not always standardized across countries and time which, among other issues, hampers comparability.

Objective

MSME-CI records the number of formally registered micro, small and medium enterprises (MSMEs) across 155 economies. This database is current as of May 2014 and expands on the August 2010 MSME-CI edition. Data is available at http://msmecountryindicators.smefinanceforum.org/msme_data_with_notes.xlsx.

The MSME-CI database attempts to provide an objective and unique overview of the MSME sector. It can be used together with other data sources. The MSME-CI fills the following knowledge gap: country-by-country structural indicators for the enterprise sector, based on firms' size for all regions of the world. The 2014 update includes the latest country-level: MSME definitions; number of enterprises; employment figures; sector distribution; and historical data. In addition, the 2014 update kick-starts the collection of data on: MSME contribution to economies; information on multiple MSME definitions (broken down by main variables); and data sources within a country. Furthermore, the 2014 update aims to differentiate the statistics provided not only at the micro, small-, medium-, and large-enterprise levels but also in aggregate results, including total small and medium enterprises (SMEs) as well as total MSMEs.

Data Collection and Scope

Data was gathered online. For cases where data was not available online, national statistical offices or other MSME-related institutions were contacted for information; however, the reply rate and information provided were very limited. In the case of European Union countries, Structural Business Statistics provided by Eurostat or SME Performance Review were used, in addition to information provided by national statistical offices. For other countries, any available information from national statistical institutes, MSME agencies, etc. was searched, evaluated and utilized. All data sources are cited.

Unless otherwise mentioned in the country-specific comments of the MSME-CI MS Excel workbook:

- Data includes enterprises without employees. This is a methodological change from the January 2010 "Methodology Note on the MSME Country Indicators (MSME-CI)."
- Data covers the private business economy. It does not cover public administration or non-market services such as health and education.
- Data generally does not include enterprises active in the agricultural sector, because most structural business statistics focus on the business economy.[1]

[1] There are several observations for the share of agriculture in the sector breakdown (33 observations for microenterprises and 43 observations for SMEs). This information was collected because it is often mentioned in descriptive sections of MSMEs studies. However, when detailed agricultural data is available, it is generally part of different data-collection processes.

Structure

The MSME-CI 2014 database 2014 displays the following main worksheets.

TABLE 1: Characteristics of the MSME-CI main worksheets

Worksheet Name	Description	Number of Countries	Number of Information Lines
Definition & Data	It contains MSME data—such as gross national income (GNI) per capita, population, number of enterprises, employees, densities, value added, etc.—as well as MSME definitions based on number of employees, assets and turnover. In this worksheet, more than one observation per country can be found.	155	267
Historical Data	It contains historical information concerning the number of enterprises, number of employees and derived variables (such as density, and share of employment broken down by microenterprise, SMEs, MSMEs, and large enterprises) according to MSME definitions based on number of employees. For several countries the information covers a span of 10 years or more.	155	1,100
Data Sources	It details the source of information for each country regarding the data presented on the Definitions & Data worksheet.	155	288

Additional descriptions for selected variables are detailed in the following subtitles. In general, the variables in each worksheet are similar and in most cases self-explanatory.

Notes on Selected Variables

Source Code

It classifies the source of information under the following criteria:

TABLE 2: Classification of source codes[2]

Source Code	Data Coming From
1	National Statistics Office
2	Central Bank, Banking Association, Ministry of Finance, Ministry of Economy, or similar[2]
3	Small Business Administration/Development/ Promotion Agencies, SME Regional Institutions, SME Associations, etc.; also, Ministry of Industry or similar

GNI Per Capita, Population, and Income Groups

Data on GNI per capita, Atlas method, and total population come from the World Bank's World Development Indicators (WDI) central database. When data is unavailable in WDI, data from the United States Central Intelligence Agency's The World Factbook is used in some cases.

October 2013 ceilings for the classification of countries by income group using GNI per capita, Atlas method, from the World Bank List of Economies were used. However, the classification of each country corresponds to the year specified in the source for the data and definition. The historical income classification is based on the World Bank Analytical Classifications, which provides information from 1987 to 2012.

If the country observation line does not display information for the columns GNI per capita and population, it signals that the entry line provides only information about the MSME definition and this MSME definition has no data available in relation to it.

Most Widely Used Definition

More than one MSME definition was listed for each country if the information was found. The "most widely used" definition—that is, the most widely used definition for a country—is only used for the MSME data analysis section and on the data visualization tool. The key factor for identifying this definition was the existence of additional data (such as number of MSMEs, number of workers in MSMEs, share of value added by MSMEs, etc.) associated with the definition. The aim is to maximize the analysis of MSME data associated with a definition. The "most widely used definition" column indicates with a "1" which is thought to be the most widely used definition for each country. A value of "0" is assigned otherwise.[3]

Data Cleaning Process

This process is also applied only to the MSME data analysis section and on the data visualization tool. Some economies had to be excluded during the data analysis to ensure more accurate results. The data cleaning process excluded economies whose information is not census data, not covering all the sectors in the economy (except for agriculture) and/or whose information was classified as an outlier after analyzing extreme values, scatter plots and post estimation of outliers (leverage, standardized and studentized residuals, among others). For instance, United Arab Emirates was excluded because data do not cover the whole country. Puerto Rico, Ghana, Iraq, Libya, Sri Lanka, Morocco, Nepal, Uganda, and Sudan were excluded because data do not cover all sectors of the economy. Nicaragua was excluded because data cover only urban areas. Ethiopia, Malawi, Mauritius, Nigeria, Guinea, and Montenegro were also excluded because data come from surveys. Finally, data for Maldives, Tanzania, Qatar, and Kuwait were excluded based on the internal analysis of outliers that was executed. Most of these cases are also highlighted in the country-specific comments in the MSME-CI MS Excel workbook.

It is important to note the information regarding the data cleaning process was included in the database in case the user would like to replicate the results of the data analysis section. However, researchers are free to choose other methodologies for the data cleaning process (determination of outliers in particular). The column "clean" indicates a value of "1" for economies whose information

[2] Those governmental institutions related to finance and economic policies are classified under source code 2 along with banking associations, central banks and others. The rest of government and non-government institutions related to MSMEs activities are classified under source code 3.

[3] Please, see Annex I for further details.

is census data and was not classified as an outlier. A value of "0" is assigned otherwise. Observations that do not provide information on MSME data also have a value of "0" in the column "clean."

MSMEs and MSMEs2

The variable MSMEs is calculated for those countries where all pieces of information are available. For example, this column was calculated if a country has information for micro, small, and medium enterprises. Alternatively, it is also calculated if the information for microenterprises and SMEs is available, or when only the information for MSMEs as a whole is available. Table 3 explains the calculation of MSMEs. The variable MSMEs attempts to capture the number of micro, small and medium enterprises in a country.

MSMEs2 attempts to capture all the available information regardless of the existence of all the pieces. For example, a country with micro and SME data contributes to MSME2, as does a country where only SME data is available. It presents the number of enterprises for all the countries who have some information available. Table 4 details the calculation of MSMEs2.

Calculations for absolute numbers in the analysis note—such as total number of MSMEs and total number of employees in MSMEs—are based on information from the MSMEs2 variable of the database. All of the other data analyses—such as MSME densities, shares of enterprises, etc.—are based on the information provided in columns micro, SMEs, and MSMEs after applying a data cleaning process.

Value Added

In many cases, the methodology used to calculate the MSMEs' contribution to an economy's value added or gross domestic product (GDP) is not specified. Major progress needs to be made to collect more and better quality information, given that this indicator is extremely important. The available information for value added for all European Union countries comes from Eurostat. In the remaining countries where value added or contribution to GDP information are available, the methodology might not be standardized (sources are mentioned in the database). Eurostat classifies value added at factor cost as an indicator in the domain of structural business statistics calculated as "the gross income from operating activities after adjusting for operating subsidies and indirect taxes." According to the United Kingdom's Office for National Statistics, gross value added (GVA) equals GDP at current market prices over all industries, plus taxes on products and minus subsidies on products.

TABLE 3: Calculation of the variable MSMEs

	Micro (available)	Small (available)	Medium (available)	SMEs (available)	MSMEs (available)	MSMEs (calculated)
Data	Yes	Yes	Yes	—	—	Yes
	Yes	—	—	Yes	—	Yes
	—	—	—	—	Yes	Yes

TABLE 4: Calculation of MSMEs2

	Micro (available)	Small (available)	Medium (available)	SMEs (available)	MSMEs (available)	MSMEs2 (calculated)
Data	Yes	—	—	—	—	Yes
	—	Yes	—	—	—	Yes
	—	—	Yes	—	—	Yes
	—	Yes	Yes	—	—	Yes
	—	—	—	Yes	—	Yes
	—	—	—	—	Yes	Yes
	Yes	Yes	Yes	—	—	Yes
	Yes	—	—	Yes	—	Yes

Sector Distribution

Only the most recent data on sector breakdown are included. If available, the sector information for MSMEs has been separated between microenterprises and SMEs. The sectors included will vary case-by-case according to the available data. The total values in the columns "Number of MSMEs" may differ between individual country worksheets and the data summarized in the "Definitions & Data" worksheet, because some sector breakdowns (detailed in the individual country worksheet) were obtained from sources different than the data featured in the worksheet "Definitions & Data." For some countries, absolute numbers of MSMEs were provided by one source, but the MSME sector breakdown was found in another source. In addition, there were cases in which the original sources did not add up in their own reported numbers. At times, it was specified that some numbers were not reported for national security or other reasons. The results of the information displayed in the individual country worksheets are summarized in the columns "Micro Enterprises: Sector Distribution" and "SME: Sector Distribution" in the "Definitions & Data" worksheet.

Other Variables

The rest of the variables in the database are self-explanatory.

Data Notes

MSME Definition

MSME-CI does not define MSME. For MSME definitions used and/or legally adopted by countries please see the database and the note "How Do Economies Define MSMEs?"[4]

In the MSME-CI, MSMEs are defined mainly by size of employment. For example, the most common thresholds are: microenterprises, less than 10 employees; small, less than 50 employees; and medium, less than 250 employees. There are some cases where a definition was not found, but data is available for several thresholds based on number of employees. For example, a national statistical office provided data for three thresholds: less than 20 employees; less than 100 employees; and equal to or more than 100 employees. Nevertheless, no official MSME definition was found in that instance. In such cases, data was structured based on the available thresholds the original source used to present the data. The results were then presented for reporting purposes. In the previous example, for instance, the data provided for 100 or more employees was reported as large; the data provided for less than 100 employees was reported as medium; and the data provided for less than 20 employees was reported as small.

Data Gaps

The following step was taken to avoid data gaps: If the data segment—such as number of MSMEs in a certain sector, GNI per capita, etc.—is not available for all the years, it is replaced with available data from the previous year or most recent year (if the previous year data are unavailable). Please see the country-specific comments of the MSME-CI Excel workbook for such occurrences.

EUROSTAT and SME Performance Review EU

Unless otherwise mentioned:

- Data from EUROSTAT's Structural Business Statistics (SBS) and SME Performance Review EU are aggregated based on type of economic activity and the size of employment.

  Sector distribution, size breakdown, number of enterprises, and number of employees follow the format and definitions of Eurostat. It includes codes B to N and S95 according to NACE Rev.2. It includes "industry, construction, and distributive trades and services. Note that financial services (NACE Rev. 2 Section K) are kept separate because of their specific nature and the limited availability of most types of standard business statistics in this area. SBS does not cover agriculture, forestry and fishing, nor public administration and (largely) non-market services such as education and health. For information on these areas of the economy, refer to national accounts by branch or other sector specific statistics."

4 "How Do Economies Define MSMEs?" (Gonzales, Mirmulstein, & Hommes, 2014).

Breakdown utilized in SBS:

- Total
- 0–9 (0 to 9 persons employed)
- 10–19 (10 to 19 persons employed)
- 20–49 (20 to 49 persons employed)
- 50–249 (50 to 249 persons employed)
- GE250 (250 or more persons employed)

Terms "Country" and "Economy"

Both terms *country* and *economy* are used in the MSME-CI. The term *country* does not imply political independence, but refers to any territory for which authorities report separate economic statistics.

Quality Control

The following controls were executed to improve the database content and ensure its accuracy:

Lessons Learned from the Data Collection

Important lessons were drawn from the MSME data while building the MSME-CI, in particular the following:

- The experience of collecting data for the MSME-CI database indicates that good quality MSME data is extremely limited, particularly in developing countries. Given several approaches for defining an MSME, institutions engaged in MSME data reporting could take a concrete first step by creating uniform standards for collecting MSME data. One approach could be to establish standard strata which aim to satisfy several definitions while at the same time enable the collection of data under clear size classes. Figure 1 provides an example of standard strata for MSME definitions utilizing number of employees. The availability of data for several standard strata across countries could be useful for producing more rigorous studies. Under the leadership of international organizations and in coordination with national statistics offices and other national/regional institutions,

TABLE 5: Quality control checks for the MSME-CI database

No.	Quality control	Description
1	Analysis of randomly chosen cases.	Countries were selected randomly to verify the existing numbers in the database (e.g. Thailand, Turkey, United Kingdom, Poland, France, etc.).
2	Determination of extremes values and analysis of outliers.	Scatter plots were generated for numbers of enterprises and population. Extreme values for densities and number of enterprises were calculated. Finally, a post-estimation of outliers' procedure was executed for microenterprises, SMEs and MSMEs densities.
3	Verification of sums of absolute numbers of enterprises (micro + SME + only MSME = Total MSMEs2). Verification of sums of absolute numbers of jobs (micro + SME + only MSME = Total MSMEs2).	Total number of enterprises and total number of employees in MSMEs were calculated. First, calculations were made considering only those observations that have all the pieces of information (micro and SMEs). The results of this procedure show lower total numbers. Second, absolute numbers were calculated considering observations with any piece of information (MSME2).
4	Verification that SMEs summation column is based on small and medium enterprises.	The column of number of SMEs was analyzed to verify its composition by small and medium enterprises. Particular cases were identified and analyzed.
5	Verify the calculation of shares (size breakdown, share of number of employees, value added).	It was verified that for countries where information was not complete (i.e. not having micro, SME and large) in terms of number of enterprises and employees, formulas should not calculate shares. Exceptions are for countries where even though not all numbers (micro, SME and large) are presented, the source does clearly specify the share of any of the groups (micro, SME or large).

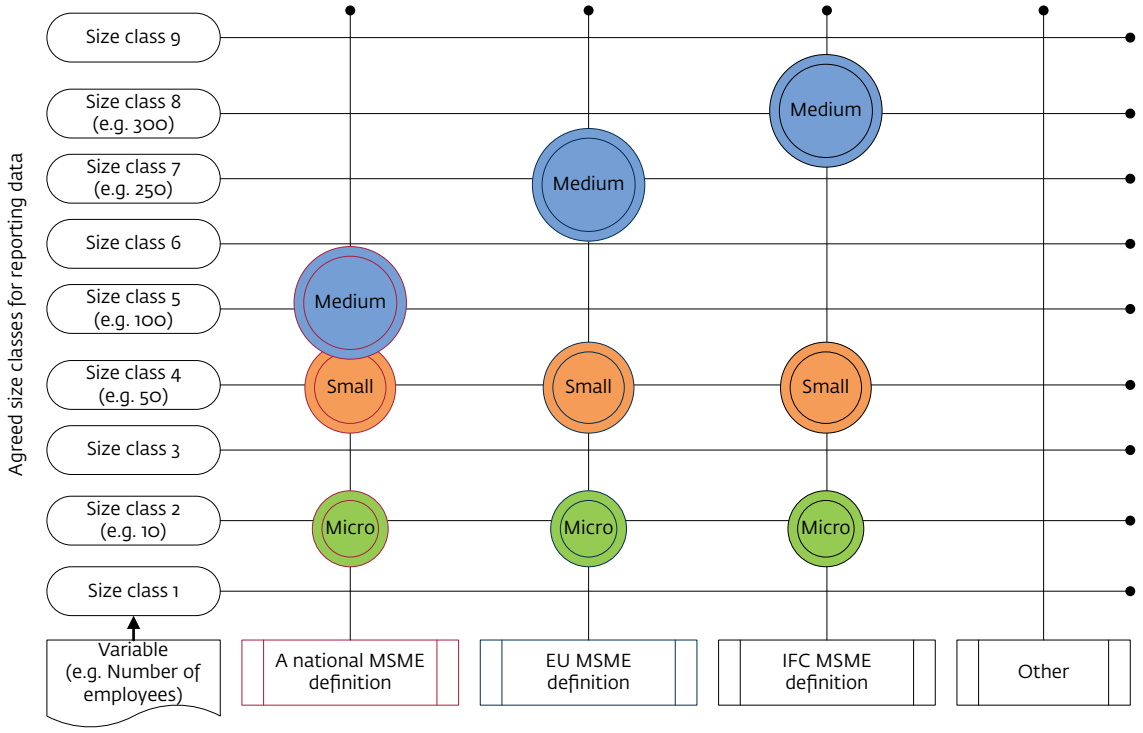
No.	Quality control	Description
6	Comparison of total number of MSMEs between the 2014 update and the 2010 update.	Calculated differences (separated by positive, negative and no difference). Positive and negative differences were prioritized by two criteria: 1) Identified the largest differences by absolute numbers 2) Identified the largest differences by percentage of variation (i.e., how large is the difference as compared with the original magnitude of the data) For negative numbers, all differences by absolute numbers above 1 million and all differences by percentage variation larger than 10 were analyzed. For positive numbers, all differences by absolute numbers above 500,000 and all differences by percentage variation larger than 100 were analyzed (increases are naturally expected).
7	In-depth verification of cases that needed further discussion.	Besides the previously described controls, some cases where the numbers provided appear to include the informal sector were also identified and discussed: for example, microenterprises in Nigeria, Indonesia, and Vietnam. Additionally, the cases of China, Azerbaijan, and Tanzania that were previously identified were selected for consultation with IFC and World Bank regional offices.
8	Compared total number of MSMEs, SMEs and microenterprises separately between the 2010 and 2014 updates.	Utilized the 2010 database to calculate the total number of MSMEs, SMEs and microenterprises. Same exercise was executed for the 2014 database (using the most widely used observation per country). Then, total numbers were compared.
9	Compared average MSME density and average MSME employment share between the 2010 and 2014 updates.	Utilized the 2010 and 2014 databases to calculate the average MSME density and average employment share. Both databases were checked to eliminate observations coming from surveys, not covering all sectors, etc. as was done in the analysis of each database.
10	Eyeball checking.	Checked tabs of the database. Verified information. Others.
11	Contrasted number of enterprises with their correspondent number of employees.	Calculated the differences for each case. Verified that number of employees for micro, SME, MSME, and large are higher than number of enterprises (or equal for enterprises with zero employees).
12	Explored the relationship of total employment in MSMEs with total employment in other databases.	Explored the relationship. Verification of sums of absolute numbers of jobs (micro + SME + only MSME =Total employment in a given country).
13	Compared total MSMEs in MSME-CI with total MSMEs in Enterprise Finance Gap Database.	Calculated differences and analyzed extreme cases. On average, MSME-CI 2014 has data that is two years more recent than the Enterprise Finance Gap Database (special focus was given to negative differences).
14	Verification of correlations between MSMEs density/MSMEs2 density with other variables (e.g. GNI per capita).	Verified if there were significant changes in the relationship of MSME density and other variables when using the variables MSMEs and MSMEs2.

MSME data collection processes could follow crowd-sourcing schemes.

- MSME data are not always standardized across countries and time. Different institutions using different methods gather data on MSMEs. These institutions define MSMEs based on differing variables and scales and sometimes change their definitions. EUROSTAT's Structural Business Statistics provides the best example of regional coordination and harmonization of MSME data.
- To have comparable MSME data, the following steps could be taken:
  - Economies should be surveyed using a unified and standardized method.
  - Institutions in charge of gathering MSME data should coordinate with each other regarding the variables and methods used to determine the size of the MSME sector.
  - Private-sector stakeholders should be engaged to help with efforts to standardize the data.

These actions can be taken first at the regional level and subsequently expanded to the global level. In return, economies would reap the benefits of a cross-country and time-series analysis of MSMEs' contribution to development.

FIGURE 1: Suggestion for harmonized collection of data: Establishment of strata for variables defining MSMEs



- Data on the informal sector was revised; but results are not reported, given that the information found is limited. Additionally, the topic may deserve more specialized attention. In general, it could be said that MSME data on the informal sector is scarce and not comparable across countries. This is due to differences in the definition of the informal sector and in estimation methods. Moreover, the informal sector seems to be defined for the whole economy without differentiating the size of the enterprises. Estimates of the informal sector are needed to make a comprehensive evaluation of the MSMEs' contribution to economic development.
- It is valuable to keep the collection of data and execute the analysis in a disaggregated way for microenterprises and SMEs. Particularly in developing nations, microenterprises seem to account for the vast majority of firms in the small enterprise sector.
- Time series data is not always available. However, it is crucial for future evaluation of the reforms of business regulations, for example.
- Some institutions collect data on MSMEs only in selected sectors, most often in manufacturing. This limits the possibilities of evaluating MSMEs' contribution to GDP or employment as a whole.
- It's vitally important to continue and improve the efforts for collecting complete data disaggregated by gender. Likewise, quality-type data such as MSMEs' share of value added, productivity, quality, and competitiveness, among others, may lead to a better understanding of MSMEs' effective contribution to the economy. Collecting these quality-type data can provide additional insights on top of what number of enterprises and number of employees could explain. The collection of data regarding number of employees working in MSMEs could also be broken down by economic sector in future updates.

Annex I: Methodological Issues

Determination of Single Values per Threshold

- When there is only a single value for each threshold within a given variable, the selection of a single value is straightforward. However, some MSME definitions may detail values differentiated by economic activity for each threshold. In the analysis, median values (if information on number of MSMEs by sector was available, weighted median values) were obtained for each threshold that was differentiated by economic sector. The table below provides an illustrative example.
- The case of China provides an interesting example supporting the use of median values. For instance, the definition of medium-sized enterprises in China establishes an upper limit of: 1,000 employees for heavy industry and transportation; 200 employees for wholesale and warehouse; 300 employees for retail, accommodation, restaurant, software, and tenancy; 2,000 employees for information transmission; and 300 employees for other industries. The median value establishes the single threshold value for medium enterprises of 300 employees. This value reduces the huge size differences among some Chinese companies and more closely resembles threshold values found in other countries across the world.

The use of weights faces a number of challenges given the limited availability of complete data. Examples where median values were weighted according to the number of MSMEs in each economic activity include Japan. The definition there specifies the three main sectors considered in their data—manufacturing, trade, and services—which allows a complete summation of all weights. However, the sector breakdown data available is for 2006, and it was applied to a definition valid for the year 2010. Other countries where median values were weighted include Thailand, Taiwan—China, Vietnam, and South Africa. By contrast, weighted figures were impossible to calculate in some countries because the definition only details thresholds for some selected economic sectors. This does not allow a complete summation of the weights. In China, for example, the definition details thresholds for construction, real estate and the commercial sector. In Kenya, the definition details thresholds for manufacturing, service and farming. Korea covers manufacturing and mining, construction, and transportation. It was also impossible to calculate weighted figures in Namibia and Togo. Furthermore, it is important to highlight that the database does not provide a sector breakdown for large enterprises.

TABLE: Example of MSME thresholds differentiated by economic activity

Country	MSME Definition				Thresholds for analysis			
	Micro	Small	Medium	Large	Micro (<)	Small (<)	Medium (<)	Large (≥)
Albania	1–4	5–19	20–79	>=80	5	20	80	80
Argentina	<5 Ind. & Trade <4 Serv.	<24 Ind. <23 Trade <17 Serv.	<96 Ind. <67 Trade <66 Serv.	<5 Ind. & Trade <4 Serv.	Median	Median	Median	Median

Source: MSME Country Indicators.

Determination of Monetary Values

- The first step was to obtain one value per cell (procedure detailed in the subtitle "Determination of Single Values per Threshold"). Subsequently, values in US$ for each threshold were calculated utilizing exchange rates[5] (local currency to US$) corresponding to each year detailed in relation to the source of data/definition information.[6] Finally, the US$ values were adjusted for inflation using the consumer price index (CPI).[7]
- Purchasing power parity (PPP) terms were not applied to the monetary figures of the database, given that they apply to GDP and private consumption levels and consequently the use of PPP for the conversion to "other Local Currency Unit series to international dollars is not recommended," according to the World Bank's Data support section.[8]

Determination of the Most Widely Used MSME Definition in a Country

- The following steps were taken for selecting the MSME definition that seem to be used the most:
  - The definition is the only one found/available in that country.
  - The definition has the most complete data associated with it.
  - The data associated with the definition represent all (or more) economic sectors.
  - The definition and data come from a national institution.
  - If, at this point, there is still more than one definition, select the latest one (year).

---

5 An important consideration is to utilize end-of-the-period figures when working with asset values, and utilize average figures when working with sales. End-of–the-year exchange rate figures were obtained for each country's corresponding data/definition year from the International Monetary Fund (IMF), and an exploratory analysis was conducted to compare these figures with the World Bank's average exchange rates that were utilized. Data from the IMF is available at http://www.imf.org/external/np/fin/data/param_rms_mth.aspx. The results of this exercise suggest that end-of-the-year figures generally present higher values than average figures, and that there is no excessive variation between average and end-of-the-year figures for our database (Libya is the only case were significant variation was found).

In addition, end-of-the-year figures provided by the IMF do not cover all the required countries and it was further analyzed that the end-of–the-fiscal-year period may differ for several countries. Not all countries report their financial figures at the end of the year. For instance, the fiscal year in Japan concludes at the end of March. Consequently, average exchange rates for each corresponding year were utilized both for assets and turnover.

6 The official exchange rates (LCU per US$, period average) from World Bank Data were utilized. They are available at http://data.worldbank.org/indicator/PA.NUS.FCRF.

7 The Consumer Price Index from the United States' Bureau of Labor Statistics (BLS) was utilized. It is available at http://www.bls.gov/cpi/data.htm.

Each observation value is adjusted with its corresponding year's CPI. Calculations were made for the original CPI provided by the BLS with the base year corresponding to the period 1982–84. In addition, values were also calculated adjusting the CPI's base year to 2011 with the intention of displaying thresholds that could be easier to interpret (the year 2011 was chosen because it is the median value of the years for the information in the 2014 update). Naturally, the results with base period 1982–84 display lower values, and no other differences where accounted for so far. A procedure detailed by the United States Census Bureau (Construction Price Indexes—Point 9) was used for changing the base year. It is available at https://www.census.gov/construction/cpi/faqs/newpriceindexqa.html#quest3.

8 Available at https://datahelpdesk.worldbank.org/knowledgebase/articles/114945-why-are-some-series-shown-in-purchasing-power-pari.

IFC International Finance Corporation
WORLD BANK GROUP

www.ifc.org