

README

Author: *Women, Business and the Law team*

This README document provides an overview of the organization and content of the **“WBL2026 DIME CodeCheck REVISED v2”** folder.

This folder contains all the materials, datasets, and scripts used to construct the WBL 2026 datasets in a fully reproducible manner.

The main directory includes four subfolders:

- 1) **docs/**: Contains documentation related to the project, including the scoring rules by pillar.
- 2) **dofiles/**: Includes all Stata scripts used to generate the WBL 2026 scores at indicator, topic and pillar levels by Economy.
- 3) **raw data/**: Stores the original input datasets prior to any processing. The WBL team utilizes a custom-built **Data Management System (DMS)** to review respondents' answers to survey questions and record verified values. “Raw data” refers to the original input datasets downloaded from the DMS following a multi-level review process to ensure accuracy for analysis, and cleaned by dropping variables irrelevant for the scoring process. The final WBL index and scores for the WBL 2026 Report were calculated using the datasets included in this folder. Please note that the raw data are not publicly available in line with the data privacy statement for WBL data collection and are not included in this Reproducibility Package. Parties interested in accessing data must contact the *Women, Business and the Law* team to request access.
- 4) **clean data/**: Contains the output datasets generated after running the do-files. Within this folder, the subfolder “clean data/temp/” holds intermediate datasets produced at different stages of the data creation process, while the main “clean data/” directory includes the final datasets for WBL 2026 data on economy level scores. Please note that the intermediate data created during the final data production are not included in this Reproducibility Package.

Data Sources:

The WBL team utilizes a custom-built **DMS** to review individual respondents' answers to survey questions and record verified values. “Raw data” refers to the original input datasets downloaded from the DMS following a multi-level review process to ensure accuracy for analysis, and cleaned by dropping variables irrelevant for the scoring process. Please note that the raw data are **not publicly available** in line with the data privacy statement for WBL data collection and are not included in this Reproducibility Package. Parties interested in accessing these datasets must contact the *Women, Business and the Law* team to request access to the following email address: wbl@worldbank.org

All files within these directories are described in detail in the sections below to facilitate transparency, replication, and efficient navigation.

Subfolder (1): “docs”	
File Name	Description
WBL 2026 EP Scoring Rules.docx	This Word document File: WBL 2026 EP Scoring Rules.docx outlines the scoring rules for the Enforcement Perceptions (EP) indicators under the <i>Women, Business and the Law</i> (WBL) 2026 cycle. It provides a detailed, step-by-step guide for calculating EP scores at the indicator, topic, and economy levels by country. The document also specifies general considerations applied in the scoring process to ensure consistency and comparability across economies.
WBL 2026 LF and SF Scoring Rules.xlsx	This Excel file File: WBL 2026 LF and SF Scoring Rules.xlsx presents the scoring rules for both the Legal Frameworks (LF) and Supportive Frameworks (SF) components under the WBL 2026 cycle. It details the scoring structure at the question, indicator, topic, and economy levels. The file includes the following key elements: WBL Topic, WBL Survey question, Data Point (DP) Hierarchy, Data Point (DP) Variable Name in Raw Data, Variable Name in Stata Coding Script, Value (Range) of Variable in coding Script, Scoring Formula, and Notes. In addition, it provides examples illustrating the estimation process for selected indicators, demonstrating how scores are calculated and aggregated across levels to ensure consistency and reproducibility.

Subfolder (2): “dofiles”	
File Name	Description

Master: 0_Master.do	This do-file serves as the central execution script for the WBL 2026 scoring process. It standardizes the Stata environment, installs all required packages, and defines global paths, used throughout the project. It structures the workflow into modular components and ensures that all subsequent task-specific do-files (e.g., for LF, SF and EP scoring) run in a controlled and reproducible manner. The third party must enter their path in order to run the dofiles and set the user to themselves.
Legal Frameworks (LF): 1_WBL_2026_LF.do	The LF do-file uses the processed WBL Legal Frameworks (LF) 2026 raw data extracted from DMS File: WBL26_DMS_LF_Cleaned_Raw_Data.dta to generate LF scores indicator by indicator across the ten topics, following the scoring rules defined in the LF Scoring Rules Excel file “WBL 2026 LF and SF Scoring Rules.xlsx”. For each topic, the do-file calculates individual indicator scores and aggregates them into topic-level scores. In the final section, it computes the overall economy-level index (LF pillar score) and applies standardized labeling to ensure consistency across outputs and saves the resulting intermediate dataset in the “clean data\temp” folder.
	- File: WBL26_LF_ALL_Scores.dta As a final step, the do-file saves the final dataset by keeping only the indicator, topic and economy index scores in the “clean data” folder: - File: WBL26_LF_Indicator_Topic_Scores_Index.dta

Supportive Frameworks (SF): 2_WBL_2026_SF.do	The SF do-file uses the processed WBL Supportive Frameworks (SF) 2026 raw data extracted from DMS - File: WBL26_DMS_SF_Cleaned_Raw_Data.dta to generate SF scores indicator by indicator across the ten topics, following the scoring rules defined in the SF Scoring Rules Excel file “WBL 2026 LF and SF Scoring Rules.xlsx”. For each topic, the do-file calculates individual indicator scores and aggregates them into topic-level scores. In the final section, it computes the overall economy-level index (SF pillar score) and applies standardized labeling to ensure consistency across outputs and saves the resulting intermediate dataset in the “clean data\temp” folder. - File: WBL26_SF_ALL_Scores.dta As a final step, the do-file saves the final dataset by keeping only the indicator, topic and economy index scores in the “clean data” folder: - File: WBL26_SF_Indicator_Topic_Scores_Index.dta
--	---

Enforcement Perceptions (EP) – Family: 3_WBL_2026_EP_Family.do	This is the 1st do-file to run for creating EP pillar scores. The EP-family do-file uses the processed WBL Enforcement Perceptions (EP) – family survey raw data - File: FAMILY_RESP_Level_EPValid.dta to estimate the EP scores indicator by indicator across 4 topics included in the Family Survey: Topic 2: Mobility, Topic 5: Marriage, Topic 8: Entrepreneurship, and Topic 9: Assets. The rules followed are detailed in the document: - File: WBL 2026 EP Scoring Rules.docx ***** Special case: For the indicator Entrepreneurship 8.4: GenderResponsive Procurement, the do-file uses an additional dataset: - File: ENT_ENF_FAM_LAB_COMBINED_RESP_Level_EPValid.dta That integrates data from the Family and Labor surveys. This portion of the do-file processes the indicator following the EP scoring rules document. The intermediate dataset with the indicator 8.4 estimated is saved in “clean data\temp” folder as: - File: Ent_Enf_GndrProc_Indicator_Score_EcoLvl.dta This intermediate dataset is later integrated to the “Entrepreneurship” topic estimation. ***** All indicators are later aggregated at topic level, and the do-file applies standardized labeling to ensure consistency across outputs. The resulting intermediate dataset is saved in the “clean data\temp” folder then to be used to create economy level final dataset for EP scores. - File: WBL26_EP_Scores_EconLevel_FAMILY.dta
--	--

Enforcement Perceptions (EP) – Labor: 4_WBL_2026_EP_Labor.do	This is the 2nd do-file to run for creating EP pillar scores. The EP-labor do-file uses the cleaned Enforcement Perceptions (EP) – labor survey data File: LABOR_RESP_Level_EPValid.dta to estimate the EP scores indicator by indicator across 5 topics included in the Labor Survey: Topic 3: Work, Topic 4: Pay, Topic 6: Parenthood, Topic 7: Childcare, Topic 10: Pension. The rules followed are detailed in the document:
	- File: WBL 2026 EP Scoring Rules.docx ***** Special case: For the indicator Work 3.1: Work Get a Job, the do file uses an additional dataset: - File: FAMILY_Wrk_GetaJob_ForLabor_RESP_Level_EPValid.dta Because this indicator was asked in the Family survey, but it belongs to the indicators in the “Work” topic that its part of Labor. This portion of the do-file processes the indicator following the EP scoring rules document. The intermediate dataset with the indicator 3.1 estimated is saved in “clean data\temp” folder as: - File: FAMILY_Wrk_GetaJob_ForLabor_ECON_Level_scores.dta This intermediate dataset is later integrated to the “Work” topic estimation. ***** All indicators are later aggregated at topic level, and the do file applies standardized labeling to ensure consistency across outputs. The resulting intermediate dataset is saved in the “clean data\temp” folder then to be used to create economy level final dataset for EP scores. - File: WBL26_EP_Scores_EconLevel_LABOR.dta

Enforcement Perceptions (EP) – VAW: 5_WBL_2026_EP_VAW.do	This is the 3rd do-file to run for creating EP pillar scores. The EP-VAW do-file uses the cleaned Enforcement Perceptions (EP) – Violence Against Women (VAW) survey data - File: VAW_RESP_Level_EPValid.dta to estimate the EP scores indicator by indicator across one topic, Topic 1: Safety. The rules followed are detailed in the document: - File: WBL 2026 EP Scoring Rules.docx All indicators are later aggregated at topic level, and the do file applies standardized labeling to ensure consistency across outputs. The resulting intermediate dataset is saved in the “clean data\temp” folder then to be used to create economy level final dataset for EP scores. - File: WBL26_EP_Scores_EconLevel_VAW.dta
Enforcement Perceptions (EP)merge: 6_WBL_2026_EP_merge.do	This is the 4th do-file to run for creating EP pillar scores. This do-file merges all datasets containing indicator, topic, and economy level EP scores produced from the previous 3 do-files (Family, Labor, and VAW surveys). Files: WBL26_EP_Scores_EconLevel_FAMILY.dta WBL26_EP_Scores_EconLevel_LABOR.dta WBL26_EP_Scores_EconLevel_VAW.dta After consolidating these datasets, the do-file generates the overall EP index score for each in the “clean data” folder and saves the final dataset: - File: WBL26_EP_Indicator_Topic_Scores_Index.dta

Subfolder (3): “raw data”	
File Name	Description
WBL26_DMS_LF_Cleaned_Raw_Data.dta	Contains the raw data required to estimate the LF scores using do-file 1_WBL_2026_LF.do . Please note that this dataset is not available for public access.

WBL26_DMS_SF_Cleaned_Raw_Data.dta	Contains the raw data required to estimate the LF scores using do-file 2_WBL_2026_SF.do . Please note that this dataset is not available for public access.
FAMILY_RESP_Level_EPValid.dta	Contains the raw data required to estimate the EPfamily survey EP scores using do-file 3_WBL_2026_EP_Family.do . Please note that this dataset is not available for public access.
ENT_ENF_FAM_LAB_COMBINED_RESP_Level_EPValid.dta	Additional dataset that combines data from the family and labor surveys needed to estimate the indicator Entrepreneurship 8.4: Gender-Responsive Procurement. This additional dataset is required to estimate the EP-family survey EP scores using do-file 3_WBL_2026_EP_Family.do . Please note that this dataset is not available for public access.
LABOR_RESP_Level_EPValid.dta	Contains the raw data required to estimate the EPlabor survey EP scores using do-file 4_WBL_2026_EP_Labor.do . Please note that this dataset is not available for public access. Please note that this dataset is not available for public access.
FAMILY_Wrk_GetaJob_ForLabor_RESP_Level_EPValid.dta	Additional dataset from the family surveys needed to estimate the indicator Work 3.1: Work Get a Job. This additional dataset is required to estimate the EPlabor survey EP scores using do-file 4_WBL_2026_EP_Labor.do . Please note that this dataset is not available for public access.
VAW_RESP_Level_EPValid.dta	Contains the raw data required to estimate the EP-VAW survey EP scores using do-file 5_WBL_2026_EP_VAW.do . Please note that this dataset is not available for public access.

Subfolder (4): “clean data”	
File Name	Description
WBL26_LF_Indicator_Topic_Scores_Index.dta	This dataset includes all WBL 2026 LF scores at indicator, topic, and economy level by economy for a total of 178 economies.

WBL26_SF_Indicator_Topic_Scores_Index.dta	This dataset includes all WBL 2026 SF scores at indicator, topic, and economy level by economy for a total of 178 economies.
WBL26_EP_Indicator_Topic_Scores_Index.dta	This dataset includes all WBL 2026 EP scores at indicator, topic, and economy level by economy for a total of 178 economies.
Subfolder “clean data/temp”	This subfolder includes the intermediate datasets created during creation data creation process. There are 7 dtafiles in this subfolder: • Ent_Enf_GndrProc_Indicator_Score_EcoLvl.dta • FAMILY_Wrk_GetaJob_ForLabor_ECON_Level_scores.dta • WBL26_EP_Scores_EconLevel_FAMILY.dta • WBL26_EP_Scores_EconLevel_LABOR.dta • WBL26_EP_Scores_EconLevel_VAW.dta • WBL26_LF_ALL_Scores.dta • WBL26_SF_ALL_Scores.dta Please note that these intermediate datasets are not available for public access and are not included in this package