

# README for Reproducibility Packages

## Project Overview

The reproducibility package reproduces the figures and tables of the paper “Carbon Border Adjustment Mechanism (CBAM) Exposure Indices – Methodological Note” by Maryla Maliszewska, Carolyn Fischer, Maksym Chepeliev and Euijin Jung. Some data files required to run the analysis are located under the folder ‘Data’. Scripts required to reproduce the results in the paper using the raw datasets are located in the categorized folders and generate outputs that are saved under the folder ‘Output’.

## Data Availability Statement & Provenance Statements

☐ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

## Statement about Rights

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

## Summary of Availability

- ☐ All data are publicly available.
- ☒ Some data cannot be made publicly available.
- ☐ No data can be made publicly available.

CO2 emission intensity data were purchased under the contract with the GTAP center. We do not have a license to disclose raw trade data obtained from the World Bank’s WITS. CO2 emission intensity data is made available exclusively for this research project through the corresponding author and the GTAP center. While at the time of the writing, we cannot share the data for other projects, we can provide access for replication purposes of this study (conditional on the signing of a confidentiality agreement and a security agreement). Individuals interested in accessing the data for replication purposes can contact Maryla Maliszewska ([Mmaliszewska@worldbank.org](mailto:Mmaliszewska@worldbank.org)). We will provide reasonable assistance with requests for clarification and replication.

## Details on each Data Source

We list the sources of raw data, where to download them, and their names. We extract the needed data for analysis from the source, cleaned them, and maintain the folder structure for easy replication. While we provide some raw data, we lack the license to publish data from the World Bank’s WITS. The ones in red are not publicly available.

## Data Sources Availability

- EMAG.CSV, EXAG2.CSV and OTAG.CSV:** **EMAG.CSV, EXAG2.CSV and OTAG.CSV:** The dataset from Chepeliev and Corong (2022) and Chepeliev et al. (2022) is available through a GTAP center contract with the World Bank and is not publicly accessible without purchase. WB staff can contact Senior Economist Maryla Maliszewska at ETIRI, [mmaliszewska@worldbank.org](mailto:mmaliszewska@worldbank.org) for access. Non-WB staff should visit <https://www.gtap.agecon.purdue.edu/databases/pricing.asp> and contact GTAP center at [contactgtap@purdue.edu](mailto:contactgtap@purdue.edu).
- WITS\_HS2017\_Bilateral\_2022.xlsx:** Query on UN COMTRADE database executed through World Integrated Trade Solution (<https://wits.worldbank.org/>). To replicate the parameters of the search, please choose HS2017 in the Nomenclature, insert the list of HS codes identified from the annex 1 of [CBAM legislation](#) in ProductCodes, select every country excluding European Union in ReporterNames, select every country in PartnerNames, select 2022 in Year and select Import in TradeFlowName. This dataset was created in 2024.
- gdp.dta:** 2022 GDP (current US\$) data is obtained from World Bank WDI in 2024. <https://databank.worldbank.org/source/world-development-indicators> Gross domestic product (current prices U.S. dollars) of Taiwan, Venezuela, Eritrea, and South Sudan data is from IMF World Economic Outlook in 2024. <https://www.imf.org/en/Publications/WEO/weo-database/2024/October/select-country-group>
- ISO\_Region\_Income.dta:** This is based on World Bank's country classification. Link is [here](#).

## Raw and processed dataset

Dataset	Description	Notes	Available to publish
WITS_HS2017_Bilateral_2022.xlsx	Raw dataset providing imports of reporting countries from partner countries in 2022, using HS2017 nomenclature.	No license	No
EMAG.CSV	Raw dataset providing greenhouse gas (GHG) emission intensity by country, sector, scope and type of GHG.	Commercially available	No
EXAG2.CSV	Raw dataset providing bilateral trade of reporting and partner countries.	Commercially available	No
OTAG.CSV	Raw dataset providing output of countries by sector.	Commercially available	No
GHG_EI_CBAM.dta	Processed dataset that provides CO2 emission intensity of aluminum, cement, fertilizer and iron and steel in 2017.	Commercially available	Yes
gdp.dta	Processed dataset at country level and USD current price in 2022	Public	Yes
ISO_Region_Income.dta	World Bank country classification in 2024	Public	Yes

CBAM_RP.xlsx	Processed dataset at country level as a result of running the code.	Public	Yes
CBAM_RP_figures.xlsx	Processed data with figures manually created.	Public	Yes

Note: files in red are not available publicly.

## Computational Requirements

### Software requirements

Required software is Stata. The code was run in version 18, but it perfectly works in any previous Stata version. The main difference between Stata 18 and older versions is the appearance of graphs. Our master do file includes a global setting that harmonizes the figure's appearances whether the code is run in version 18 or any previous one.

### Memory and runtime requirements

The code was run for the last time in a Dell laptop with Windows 11 Enterprise version, 16 GB RAM, and an 11<sup>th</sup> Gen Intel ® Core™ processor. A replicator could expect the whole code to run in less than 10 minutes.

## Instructions for Replicators

### Folder structure

For full replication of the project, we suggest the following, complete folder structure.

- 1\_Data
- 2\_Code
- 3\_Output

The code in the folder 'Code' will reproduce some of results included in the report. Please download codes from folder 'Code' and raw data from 'Data' and follow the steps.

- Open the folder and navigate to codes. Please update the directory to your own, following the code's direction.
- Run a do file to generate the underlying data to create the figures.
- All figures in the manuscript are manually created.

### List of Exhibits

The provided code and data reproduce:

- ☐ All numbers provided in text in the paper
- ☒ All tables and figures in the paper

- Selected tables and figures in the paper, as explained and justified below

Exhibit file name	Output filename	Code
Figure 1.a and 1.b	RP_figures.xlsx/fig1 g_alu_t.png; g_alu_t_sc.png; g_cem_t.png; g_cem_t_sc.png; g_fert_t.png; g_fert_t_sc.png; g_iron_t.png; g_iron_t_sc.png;	Line 267- 322
Figure 2	RP_figures.xlsx/fig2 g_atel.png	Line 323-331
Figure 3.a and 3.b	RP_figures.xlsx/fig3 g_alu_s.png; g_alu_s_sc.png; g_cem_s.png; g_cem_s_sc.png; g_fert_s.png; g_fert_s_sc.png; g_iron_s.png; g_iron_s_sc.png;	Line 335-375
Figure 4	RP_figures.xlsx/fig4 g_eei.png	Line 382-392

## Description of programs and code

The project is organized as follows. First, we use the raw data files to create the datasets used in the analysis. The replicator can adjust the links in these files centrally without needing to adjust the (relative) paths in the specific cleaning or analysis parts.

### Creation of data sets for analysis

- cbam\_rp.do

This do-file takes raw and processed datasets and calculates the CBAM exposure index.

Inputs:

\$main\Data\gdp.dta

\$main\Data\GHG\_EI\_CBAM.dta

\$main\Data\ISO\_Region\_Income.dta

\$main\Data\WITS\_HS2017\_Bilateral\_2022.xlsx

\$main\Data\OTAG.CSV

\$main\Data\EMAG.CSV

\$main\Data\EXAG2.CSV

Outputs:

\$main\Output\CBAM\_RP.xlsx

\$main\Output\CBAM\_RP\_figures.xlsx

\$main\Output\g\_alu\_t.png;

\$main\Output\g\_alu\_t\_sc.png;

\$main\Output\g\_cem\_t.png;

\$main\Output\g\_cem\_t\_sc.png;

\$main\Output\g\_fert\_t.png;

\$main\Output\g\_fert\_t\_sc.png;

\$main\Output\g\_iron\_t.png;

\$main\Output\g\_iron\_t\_sc.png;

\$main\Output\g\_atei.png

\$main\Output\g\_alu\_s.png;

\$main\Output\g\_alu\_s\_sc.png;

\$main\Output\g\_cem\_s.png;

\$main\Output\g\_cem\_s\_sc.png;

\$main\Output\g\_fert\_s.png;

\$main\Output\g\_fert\_s\_sc.png;

\$main\Output\g\_iron\_s.png;

\$main\Output\g\_iron\_s\_sc.png;

\$main\Output\g\_eei.png