

# README for the Reproducibility Package for Due diligence in supply chains: the case of the EUDR

## Project Overview

The reproducibility package reproduces the figures of the paper “Due diligence in supply chains: the case of the EUDR” by Shane Sela, Melvin Spreij and Iryna Sikora. 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.

## 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 raw data, we lack the license to publish all files. The ones in red are not publicly available.

## Data Sources Availability

- **2018, 2019, 2020, 2021, 2022 and 2023.csv in Data folder:** Query on UN COMTRADE database executed through World Integrated Trade Solution (<https://wits.worldbank.org/WITS/WITS/AdvanceQuery/RawTradeData/QueryDefinition.aspx?Page=RawTradeData>). To replicate the parameters of the search, please choose HS2017 in the Nomenclature, insert the list of HS codes identified from the ‘HS codes for policy product.xlsx’ in Data folder for ProductCodes, select every country excluding European Union in ReporterNames, select every country in PartnerNames, select 2018, 2019, 2020, 2021, 2022 or 2023 in Year and select Import in TradeFlowName.

- **2018e, 2019e, 2020e, 2021e, 2022e and 2023e.csv in Data folder:** Query on UN COMTRADE database executed through World Integrated Trade Solution (<https://wits.worldbank.org/WITS/WITS/AdvanceQuery/RawTradeData/QueryDefinition.aspx?Page=RawTradeData>). To replicate the parameters of the search, please choose HS2017 in the Nomenclature, insert the list of HS codes identified from the 'HS codes for policy product.xlsx' in Data folder for ProductCodes, select every country excluding European Union in ReporterNames, select every country in PartnerNames, select 2018, 2019, 2020, 2021, 2022 or 2023 in Year and select Export in TradeFlowName.
- **'GDP 2018-2023.xlsx' in Data folder:** GDP (current US\$) data is obtained from World Bank WDI in 2025. <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' in Data folder:** This is based on World Bank's country classification. Link is [here](#).
- **TME.csv in Data folder:** Query on UN COMTRADE database executed through World Integrated Trade Solution (<https://wits.worldbank.org/WITS/WITS/AdvanceQuery/RawTradeData/QueryDefinition.aspx?Page=RawTradeData>). To replicate the parameters of the search, please choose HS2017 in the Nomenclature, select all products for ProductCodes, select every country in ReporterNames, select every country in PartnerNames, select 2018, 2019, 2020, 2021, 2022 or 2023 in Year and select Export and Import in TradeFlowName.

### Data Sources Availability

'Data' folder	Filename	Description	Notes
Data folder	2018, 2019, 2020, 2021, 2022 and 2023.csv	Bilateral trade flows (imports) for selected products	Not included in the reproducibility package due to no redistribution license
Data folder	2018e, 2019e, 2020e, 2021e, 2022e and 2023e.csv	Bilateral trade flows (exports) for selected products	Not included in the reproducibility package due to no redistribution license
Data folder	GDP 2018-2023.csv	GDP data for all countries	Public
Data folder	ISO_Region_Income.dta	World Bank country classification in 2024	Public
Data folder	TME.csv	Bilateral trade flows (imports and exports) for all products	Not included in the reproducibility package due to no

Note: files in red are available publicly but not included in the reproducibility package.

## Computational Requirements

### Software requirements

Required software is Stata. The code was run in version 18, but it perfectly works in any previous or posterior 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 note. 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

Table below provides a mapping between all figures of the note and the codes producing these results.

Exhibit name	Output filename	Code	Notes
Figure 1	Figure1.png	eudr_note.do (line 1-14)	Figure shows as output of the code
Figure 2	Figure2.png	eudr_note.do (line 16-18)	Figure shows as output of the code
Figure 3	Figure3.png	eudr_note.do (line 20-22)	Figure shows as output of the code
Figure 4	Figure4.png	eudr_note.do (line 24-26)	Figure shows as output of the code

### Description of programs and code

The project is organized as follows. Main do file 'eudr\_note.do' calls for other do files to perform tasks.

- datasets.do – makes the main dataset from raw data (used for this and other projects)
- dataset\_eudr\_2022\_regions.do – modifies the main dataset focusing on EUDR products in year 2022 only and adds region averages
- dataset\_comm\_2022\_countries.do – modifies the main dataset focusing on EUDR products in year 2022 only
- figure\*.do files – create the four figures used in the note.