Assessing nature-related risks in the Hungarian financial system

Data Analysis Handbook

Version 1.1

30/05/2024

# Purpose, scope, and use of the Handbook

This handbook is part of the data analysis package provided as a complement to the OECD Environment Working Paper titled “Assessing Nature-Related Risks in the Hungarian Financial System: Charting the Impact of Nature's Financial Echo”. It aims to provide a comprehensive guide to the data analysis tools and procedures used in the paper and serves as a reference for replicating the study in the future with updated data. Additionally, it aims to facilitate the understanding and transparency of the analytical processes, ensuring that the methods and results can be accurately reproduced and validated by other researchers and stakeholders.

This handbook is intended for researchers, analysts, and policymakers involved in assessing nature-related financial risks within the Hungarian financial system or similar contexts. It outlines the folder structure used to organize data, code, documentation, and outputs, facilitating easy navigation and management of project files. Furthermore, it includes instructions for replicating the study annually with updated data. Ensuring transparency and reproducibility, this handbook is designed to be a practical and user-friendly resource. Nonetheless, some elementary coding knowledge on Stata would be particularly beneficial for more advanced studies using this tool, enabling users to efficiently conduct similar analyses and contribute to the ongoing assessment and management of nature-related financial risks.

For guidance on interpreting the results generated from the analysis please refer to the Working Paper, and for an in-depth explanation of the data cleaning, processing, and analysis methods employed please refer to its Annex, as these will not be covered within this Handbook.

# Data sources and future replication guidelines

The data analysis was conducted using Stata/MP 18.0 on Microsoft Windows Server 2019 Standard. Users on macOS should be wary about file paths included in the do-files, particularly relating to the use of backslashes and forward slashes.

The study uses data from 4 main sources for the quantitative aspects of the analysis. All data files are stored in the “0\_rawdata” folder. The data package, in the form it is shared with MNB, do not include any confidential information. Therefore, in order to replicate the study, the two .sas7bdat files that were shared by the MNB on “Instruments” and “Collaterals” must be placed in the “0\_rawdata\MNB” folder with these same exact names.

For future replication of the study, the new versions of these data files can be replaced in the “0\_rawdata” folder, making sure to respect the original file names, file formats, and data formats. Even though special attention was paid to future-proofing and easing the replication of the study with new data, the complexity of the analysis and the fact that data from many different sources were combined for the study inevitably makes this prone to requiring the modification of certain parts of the code if a data structure changes significantly from one year to the next.

1. MNB

MNB has provided data on instruments held by non-financial (non-banking, to be specific) corporates and individuals classified as primary agricultural producers at the end of 2022, as well as on the associated collaterals. Particular attention must be paid to extracting any future instruments and collaterals data from the Bank’s administrative systems in the same format as the original data shared with the OECD in late 2023.

1. ENCORE

The impact and dependency materiality ratings were provided privately by UNEP-WCMC. A publicly available version of the data can be found on the [ENCORE website](https://encorenature.org/en). The ENCORE database is expected to undergo a significant overhaul soon, which, despite being great news for assessing nature and biodiversity related risks, might make parts of the assessment obsolete. In that case, the do-files “2\_nace\_impacts\_dependencies” and “3\_exiobase\_impacts\_dependencies” would need to be modified to incorporate the new set of materiality ratings in the rest of the analysis.

1. EXIOBASE

[EXIOBASE 3](https://zenodo.org/records/5589597), the input-output data chosen for this study, is hosted on Zenodo, and can be downloaded for the relevant year of study.

1. FIGARO

[European Commission’s FIGARO data on inter-country supply, use, and input-tables](https://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/database#CSV%20flat%20format%20(FIGARO%202023%20edition)) are hosted on [Eurostat](https://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/database#CSV%20flat%20format%20(FIGARO%202023%20edition)). Having been compiled using Nace classification of economic activities, FIGARO data was particularly useful in some applications such as exploring summary statistics of Hungarian industries, however due to the lack of sufficient granularity in the breakdown of these economic activities, was not used as the basis for the IO model.

1. Auxiliary

There are also a range of auxiliary data files that are required for the code to run smoothly. The OECD Exchange rates (“OECD\_exchange\_rates”) of the analysis year should be downloaded from the OECD Data Explorer. (The initial implementation of the code uses data downloaded from OECD’s iLibrary. OECD data will no longer be hosted on the iLibrary as of 2023, and will be moved to the Data Explorer. Unfortunately, this might cause a change in the structure of the data, requiring a modification of the sections of the code on standardizing different currencies within “1\_mnb\_cleaning”. The remaining files, including “nace\_rev2.dta” and “nace\_rev2\_1\_4.dta” which contain the Nace industry classifications, “iso\_country\_codes” which contain names and ISO codes of various countries, “ins\_type\_HM\_Groupings” which classifies the various instruments found in the Hungarian data into overarching categories, and “CPRS\_20220909\_NGFS” which contains information on CPRS classifications of Nace codes, will not require regular updates. Finally, the folder “Map\_data” has the relevant Hungarian shapefile data required to construct the maps.

# Folder structure and required files

The following outlines the required structure of the "data\_analysis" folder, listing the necessary folders and files to run the "0\_master" do-file. Only the essential files and folders are included; intermediary files generated by the do-files are not listed. File names are indicated in quotes, while folder names are not.

1. data\_analysis
   1. 0\_rawdata
      1. MNB

“Instruments.sas7bdat”

“Collaterals.sas7bdat”

“data\_dictionary.xlsx”

* + 1. ENCORE

“impact\_materialities.xlsx”

“dependency\_materialities.csv”

“nace\_productionprocess\_corrtable.dta”

* + 1. EXIOBASE

IOT\_2022\_ixi

“NACE2full\_EXIOBASEp.xlsx”

* + 1. FIGARO

“matrix\_eu-ic-io\_ind-by-ind\_23ed\_2021”

* + 1. Auxiliary

Map\_data

“CPRS\_20220909\_NGFS.xlsx”

“ins\_type\_HM\_Groupings.xlsx”

“iso\_country\_codes.csv”

“nace\_rev2.dta”

“nace\_rev2\_1\_4.dta”

“OECD\_exchange\_rates.csv”

* 1. 1\_code

“0\_master.do”

“1\_mnb\_cleaning.do”

“2\_nace\_impacts\_dependencies.do”

“3\_exiobase\_impacts\_dependencies.do”

“4\_exiobase\_cleaning.do”

“5\_exiobase\_cleaning\_sector.do”

“6\_figaro\_cleaning.do”

“7\_identification\_and\_prioritisation.do”

“8\_economic\_risk\_assessment.do”

“9\_financial\_risk\_assessment.do”

* 1. 2\_documentation
  2. 3\_processeddata
     1. MNB
     2. ENCORE
     3. EXIOBASE
        1. exiobase\_impacts\_dependencies
     4. FIGARO
  3. 4\_output
     1. Identification\_and\_prioritisation
     2. Economic\_risk\_assessment
     3. Financial\_risk\_assessment

# Code summary, inputs, and outputs

### 0\_master

This do-file is the master do-file for the project on the implementation of the Supervisory Framework for Assessing Nature-related Financial Risks for the Hungarian financial system. The master do-file is the only do-file that needs to be run in order to replicate the study as it has been published; it runs the other do-files itself.

The master do-file has three sections: 1) Set directories, 2) Select process, and 3) Execution.

In section 1) Set directories, a new user would need to replace the directory pathway with information of their own computer, or add it to the end of the section and comment out the previous users’ information. More specifically, the *if c(username) == ““* line needs the username of the computer profile to be inserted in between quotes. A user can identify its username by typing *di c(username)* on the Stata terminal. Then, the local pathway of the data analysis folder should be inserted in between quotes on the line *global path = ““*. This will allow the master do-file to refer to the correct folders when executing the remaining do-files.

In section 2) Select process, a user can select which processes they would like to run. The processes (do-files) usually run on dependencies created by the previous processes, and therefore it is recommended to run all processes, at least the first time. However, given that certain operations, like cleaning the MNB data, might be computationally intensive, a user might then prefer to run only parts of the code in order to generate the outputs they are looking for. In this case, the user can set the local macros of the processes that they would like to run to 1, all the while setting the local macros of the processes they would like to ignore to 0.

### 1\_mnb\_cleaning.do

This do-file cleans the “instruments” and “collaterals” data sent by the MNB.

1. Inputs
   1. Main
      1. “$rawdata\MNB\Instruments.sas7bdat”
      2. “$rawdata\MNB\Collaterals.sas7bdat”
      3. “$rawdata\MNB\data\_dictionary”
   2. Auxiliary
      1. “$rawdata\Auxiliary\iso\_country\_codes.csv”
      2. “$rawdata\Auxiliary\OECD\_exchange\_rates”
      3. “$rawdata\Auxiliary\nace\_rev2\_1\_4.dta”
2. Outputs
   1. Data
      1. “$processed\MNB\instruments.dta”
      2. “$processed\MNB\instruments\_firms.dta”
      3. “$processed\MNB\collaterals.dta”

### 2\_nace\_impacts\_dependencies.do

This do-file matches the ENCORE impact and dependency materiality ratings to Nace codes.

1. Inputs
   1. Main
      1. “$rawdata\ENCORE\nace\_productionprocess\_corrtable”
      2. “$rawdata\ENCORE\impact\_materialities.xlsx”
      3. “$rawdata\ENCORE\dependency\_materialities.csv”
2. Outputs
   1. Data
      1. “$processed\ENCORE\nace\_impacts.dta”
      2. “$processed\ENCORE\nace\_dependencies.dta”
      3. “$processed\ENCORE\nace\_impacts\_high.dta”
      4. “$processed\ENCORE\nace\_dependencies\_high.dta”

### 3\_exiobase\_impacts\_dependencies.do

This do-file matches the ENCORE impact and dependency materiality ratings to Exiobase industries.

1. Inputs
   1. Main
      1. “$rawdata\Exiobase\NACE2full\_EXIOBASEp”
      2. “$rawdata\ENCORE\nace\_productionprocess\_corrtable.dta”
      3. “$rawdata\ENCORE\impact\_materialities.xlsx”
      4. “$rawdata\ENCORE\dependency\_materialities.csv”
      5. “$rawdata\Exiobase\IOT\_2022\_ixi\industries.txt”
2. Outputs
   1. Data
      1. “$processed\Exiobase\exiobase\_nace\_corrtable.dta”
      2. “$processed\Exiobase\exiobase\_encore\_corrtable.dta”
      3. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts.dta”
      4. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies.dta”
      5. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts\_high.dta”
      6. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies\_high.dta”

### 4\_exiobase\_cleaning.do

This do file cleans the Exiobase files at the country-sector level.

1. Inputs
   1. Main
      1. “$rawdata\Exiobase\IOT\_2022\_ixi\A.txt”
      2. “$rawdata\Exiobase\IOT\_2022\_ixi\Z.txt”
      3. “$rawdata\Exiobase\IOT\_2022\_ixi\X.txt”
      4. “$rawdata\Exiobase\IOT\_2022\_ixi\impacts\F.txt”
      5. “$rawdata\Exiobase\IOT\_2022\_ixi\Y.txt”
2. Outputs
   1. Data
      1. Main
         1. “$processed\Exiobase\L\_2022”
         2. “$processed\Exiobase\G\_2022”
      2. Auxiliary
         1. “$processed\Exiobase\A\_2022.dta”
         2. “$processed\Exiobase\Exiobase\_industry\_list”
         3. “$processed\Exiobase\Z\_2022”
         4. “$processed\Exiobase\X\_2022”
         5. “$processed\Exiobase\diagXinv\_2022”
         6. “$processed\Exiobase\F\_2022”
         7. “$processed\Exiobase\Y\_2022”

### 5\_exiobase\_cleaning\_sector.do

This do-file cleans Exiobase data for the year 2022 at the sector level, calculates indirect (and total) impact and dependency scores, and merges them in with the MNB data.

1. Inputs
   1. Main
      1. “$rawdata\Exiobase\IOT\_2022\_ixi\X.txt”
      2. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts.dta”
      3. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies.dta”
      4. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts\_high.dta”
      5. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies\_high.dta”
      6. “$processed\MNB\instruments\_firms.dta”
   2. Auxiliary
      1. “$processed\Exiobase\exiobase\_nace\_corrtable.dta”
2. Outputs
   1. Data
      1. Main
         1. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_standard\_materialities.dta”
         2. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_cautious\_materialities.dta”
         3. “$processed\Exiobase\exiobase\_impacts\_dependencies\nace\_standard\_materialities.dta”
         4. “$processed\Exiobase\exiobase\_impacts\_dependencies\nace\_cautious\_materialities.dta”
         5. “$processed\MNB\nace\_standard\_materialities\_with\_mnb.dta”
         6. “$processed\MNB\nace\_cautious\_materialities\_with\_mnb.dta”
      2. Auxiliary
         1. “$processed\Exiobase\X\_2022”
         2. “$processed\Exiobase\Z\_2022\_sector”
         3. “$processed\Exiobase\L\_2022\_sector”
         4. “$processed\Exiobase\G\_2022\_sector”
         5. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts\_indirect.dta”
         6. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies\_indirect.dta”
         7. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts\_high\_indirect.dta”
         8. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies\_high\_indirect.dta”

### 6\_figaro\_cleaning

This do-file cleans the Figaro data and outputs the Nace 1 level IO matrix, as well as the associated summary statistics.

1. Inputs
   1. Main
      1. “$rawdata\FIGARO\matrix\_eu-ic-io\_ind-by-ind\_23ed\_2021.csv”
      2. “$processed\MNB\instruments\_firms.dta”
   2. Auxiliary
      1. “$processed\ENCORE\nace\_impacts.dta”
2. Outputs
   1. Data
      1. Main
         1. “$processed\FIGARO\X\_2021”
         2. “$processed\FIGARO\X\_2021\_nace1”
         3. “$processed\FIGARO\Z\_2021”
         4. “$processed\FIGARO\Z\_2021\_nace1”
         5. “$processed\FIGARO\L\_2021\_nace1”
         6. “$processed\FIGARO\G\_2021\_nace1”
         7. “$processed\FIGARO\summary\_statistics”

### 7\_identification\_and\_prioritisation.do

This do-file generates the outputs to be included in the “Identification and Prioritisation” section of the paper.

1. Inputs
   1. Main
      1. “$processed\MNB\nace\_standard\_materialities\_with\_mnb.dta”
      2. “$processed\MNB\nace\_cautious\_materialities\_with\_mnb.dta”
      3. “$processed\Exiobase\exiobase\_impacts\_dependencies\nace\_standard\_materialities.dta”
      4. “$processed\FIGARO\G\_2021\_nace1”
      5. “$processed\FIGARO\L\_2021\_nace1”
      6. “$processed\FIGARO\summary\_statistics”
      7. “$processed\MNB\instruments\_firms.dta”
   2. Auxiliary
      1. “$rawdata\CPRS\_NGFS\CPRS\_20220909\_NGFS”
      2. “$rawdata\Map\_data\NUTS\_RG\_01M\_2021\_3035\world.dta”
      3. “$rawdata\ Map\_data\NUTS\_RG\_01M\_2021\_3035\world\_shp.dta”
      4. “$rawdata\Map\_data\Hungary\_shapefile\hun\_admbnda\_osm\_20220720\_shp\hungary2.dta”
      5. “$rawdata\Map\_data\Hungary\_shapefile\hun\_admbnda\_osm\_20220720\_shp\hungary3\_shp.dta”
2. Outputs
   1. Exposure to Physical and Transition Risks
      1. “$output\Identification\_and\_prioritisation\portfolio\_share\_impacts\_dependencies”
   2. Impact and Dependency Links
      1. “$output\Identification\_and\_prioritisation\sankey.xlsx”
      2. “$output\Identification\_and\_prioritisation\Sankey\_impacts.emf”
      3. “$output\Identification\_and\_prioritisation\Sankey\_dependencies.emf”
   3. Portfolio scores
      1. “$output\Identification\_and\_prioritisation\ecosystem\_barchart\_high”
   4. Sectoral disaggregation
      1. “$output\nace\_impacts\_dependencies\sectoral\_breakdown\_high\_nace1”
   5. Climate and nature nexus
      1. “$output\Identification\_and\_prioritisation\hvh\_cprs\_link”
   6. Directness
      1. “$output\Identification\_and\_prioritisation\directness”
      2. “$output\Identification\_and\_prioritisation\directness\_together”
   7. Upstreamness & Downstreamness
      1. “$output\Identification\_and\_prioritisation\upstreamness”
   8. Summary Statistics
      1. “$output\Identification\_and\_prioritisation\industry\_stats”
   9. Geographical scope
      1. “$output\Identification\_and\_prioritisation\total\_exposure.emf”
      2. “$output\Identification\_and\_prioritisation\total\_exposure\_a.emf”
      3. “$output\Identification\_and\_prioritisation\total\_exposure\_c.emf”
      4. “$output\Identification\_and\_prioritisation\total\_exposure\_l.emf”
      5. “$processed\MNB\instruments\_firms\_districts”
      6. “$output\Identification\_and\_prioritisation\total\_exposure\_district.emf”
      7. “$output\Identification\_and\_prioritisation\total\_exposure\_a\_district.emf”
      8. “$output\Identification\_and\_prioritisation\total\_exposure\_c\_district.emf”
      9. “$output\Identification\_and\_prioritisation\total\_exposure\_l\_district.emf”
      10. “$output\Identification\_and\_prioritisation\top\_10\_borrower\_locations”
      11. “$output\Identification\_and\_prioritisation\top\_10\_borrower\_locations\_agriculture”

### 8\_economic\_risk\_assessment

This do-file generates the outputs to be included in the “Economic Risk Assessment” section of the paper.

1. Inputs
   1. Main
      1. “$rawdata\Exiobase\IOT\_2022\_ixi\impacts\F.txt”
      2. “$processed\Exiobase\X\_2022.dta”
      3. “$processed\Exiobase\L\_2022.dta”
      4. “$processed\Exiobase\F\_2022.dta”
      5. “$processed\Exiobase\G\_2022.dta”
      6. “$processed\Exiobase\Z\_2022.dta”
      7. “$processed\Exiobase\Y\_2022.dta”
      8. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_impacts.dta”
      9. “$processed\Exiobase\exiobase\_impacts\_dependencies\exiobase\_dependencies.dta”
   2. Auxiliary
      1. “$processed\Exiobase\exiobase\_nace\_corrtable.dta”
      2. “$processed\Exiobase\Exiobase\_industry\_list”
      3. “$rawdata\Auxiliary\nace\_rev2.dta”
2. Outputs
   1. “$output\Economic\_risk\_assessment\drought\_shock\_vector.dta”
   2. “$output\Economic\_risk\_assessment\Domestic\_scenario\_results\_Exiobase”
   3. “$output\Economic\_risk\_assessment\Foreign\_exposure\_results\_Exiobase”
   4. “$output\Economic\_risk\_assessment\Trade\_partners\_Exiobase”
   5. “$output\Economic\_risk\_assessment\Foreign\_transition\_risk\_Exiobase”
   6. “$output\Economic\_risk\_assessment\domestic\_scenario\_analysis\_nace”
   7. “$output\Economic\_risk\_assessment\domestic\_scenario\_analysis\_nace\_varn”
   8. “$output\Economic\_risk\_assessment\Foreign\_transition\_risk\_nace”
   9. “$output\Economic\_risk\_assessment\foreign\_exposure\_analysis\_nace\_table.xlsx”

### 9\_financial\_risk\_assessment

This do-file generates the outputs to be included in the “Financial Risk Assessment” section of the paper.

1. Inputs
   1. Main
      1. “$processed\MNB\instruments\_firms.dta”
   2. Auxiliary
      1. “$rawdata\Auxiliary\ins\_type\_HM\_Groupings.xlsx”
2. Outputs
   1. “$output\Financial\_risk\_assessment\instrument\_category\_by\_year”
   2. “$output\Financial\_risk\_assessment\interest\_rate\_by\_sector”
   3. “$output\Financial\_risk\_assessment\outstanding\_debt\_by\_currency”