

README.md

Replication Package

Paper Title: *Who Gains from Cybersecurity Preparedness?*

This folder contains the data files and standalone R code required to reproduce the empirical results in the paper. A third party can run this package.

1. Package Contents

Main replication script

- `WhoGains\_clean\_code\_final.R`

Input datasets

- `sectors\_digital\_content.dta`
- `SectoralEnergyIndexInput\_US.xlsx`
- `SectoralLaborInputIndex\_US.xlsx`
- `gci\_panel.xlsx`
- `Final\_Maryland.xlsx`
- `UNSD\_industry\_added\_value\_USD\_constant\_2015.xlsx`
- `us\_cyberincidents.csv`
- `classy.csv`
- `2017gci.csv`

Output folders

- `outputs/`
- `latex/`

2. Software Requirements

Use R version 4.2 or later.

Install required packages:

```
install.packages(c(
  "dplyr", "tidyr", "readr", "readxl", "haven",
  "stringr", "fixest", "sandwich", "lmtest", "writexl"
))
```

3. How to Run

Set the working directory to this folder and run:

```
source("WhoGains_clean_code_final.R")
```

Outputs will be created automatically in:

```
outputs/tables/
```

4. Data Availability Statement

`sectors\_digital\_content.dta`

This file contains sector-level digital input exposure measures used to construct alternative vulnerability proxies. It is an author-prepared file based on international input-output data sources such as OECD input-output.

Citation: OECD Input-Output, author compilation.

URL / Access: <https://www.oecd.org/en/data/datasets/inter-country-input-output-tables.html>

Access Date: 12/10/2025.

Download Steps: Visit OECD statistics portal, locate input-output tables, download sector-level data.

License / Restrictions: Open source.

`SectoralEnergyIndexInput\_US.xlsx`

This file contains U.S. sectoral energy input indices used for placebo regressions. The name of the data is "Quantity Index Energy Inputs (2017=100)" produced by the BEA.

Citation: U.S. Bureau of Economic Analysis (BEA), Industry Accounts.

URL / Access: <https://www.bea.gov/data/special-topics/integrated-industry-level-production-account-klems>

Access Date: 12/10/2025.

Download Steps: Visit BEA website URL provided → Production Account Tables, 1997–2023 XLSX → open Excel file → That workbook contains the sheet: Quantity Index Energy Inputs (2017=100) → copy/paste that worksheet into a new Excel file.

License / Restrictions: U.S. government public-use data.

`gci\_panel.xlsx`

This file contains panel data for the Global Cybersecurity Index across countries for 2017, 2018, 2020, and 2024.

ITU often does not provide a clean downloadable Excel/CSV panel file directly on its site. Instead, the country scores are in: reports. The World Bank data portal provides consolidated data for 2020 and 2024. Therefore, the dataset was built manually merging the WB data for 2020 and 2024, with the data from ITU reports for 2017 and 2018.

Citation: International Telecommunication Union (ITU), Global Cybersecurity Index; World Bank.

URL / Access 1 (World Bank consolidated): https://data360files.worldbank.org/data360-data/data/ITU_GCI/ITU_GCI_GCI_OVRL_SCRE_WIDEF.csv?utm_source=chatgpt.com

URL / Access 1 (ITU reports for 2017): https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2017-R1-PDF-E.pdf

URL / Access 1 (ITU reports for 2018): https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2018-PDF-E.pdf

Access Date: 10/01/2025.

Download Steps: Download World Bank GCI consolidated containing data for 2020 and 2024 → open Excel file → copy/paste data in new Excel file → Download ITU reports "2017 Global Cybersecurity Index" and "2018 Global Cybersecurity Index" → The data for 2017 is available in Annex 1 of the report → Multiply 2017 data by 100 and add to Excel file containing the 2020 and 2024 data → do the same for 2018 data

License / Restrictions: Open source.

`2017gci.csv`

This file contains the 2017 cross-section of the Global Cybersecurity Index. It is just a subset of the dataset "gci_panel.xlsx" done in Excel.

Citation: International Telecommunication Union (ITU), Global Cybersecurity Index 2017.

URL / Access: ITU official publications: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2017-R1-PDF-E.pdf

Access Date: 10/01/2025.

Download Steps: Download 2017 GCI report and use appendix table to build dataset.

License / Restrictions: Open source.

`Final\_Maryland.xlsx`

This file contains the global cyber incident data used to construct sectoral vulnerability measures.

Citation: University of Maryland cyber incident/event data, author-prepared extraction.

URL / Access: <https://cissm.umd.edu/cyber-events-database>

Access Date:* 01/02/2025.

Download Steps: Included in package. Access could be received by accessing the link, subscribing, and downloading.

Note: This endpoint returned cyber incidents collected from curated websites and blogs. It now returns cyber incidents using the Global Database on Events Language and Tone, which results in a dataset that could differ slightly to the data downloaded in early 2025.

License / Restrictions: Open source.

`us\_cyberincidents.csv`

This file contains U.S.-specific cyber incidents used to construct alternative vulnerability measures. This is a subset of "Final_Maryland.xlsx" constructed in Excel.

Citation: University of Maryland cyber incident/event data, author-prepared extraction.

URL / Access: Included in package. <https://cissm.umd.edu/cyber-events-database>

Access Date: 10/01/2025.

Download Steps: Included in package. Access could be received by accessing the link, subscribing, and downloading. Sorting. Creating USA dataset only.

License / Restrictions: Open source.

`UNSD\_industry\_added\_value\_USD\_constant\_2015.xlsx`

This file contains country-sector real value added data in constant 2015 USD used to compute sector growth outcomes.

Citation: United Nations Statistics Division (UNSD), National Accounts / Industry Value Added data.

URL / Access: UNData portal: <http://data.un.org/Explorer.aspx>

Access Date: 05/10/2025

Download Steps: Click on link → National Accounts Estimates of Main Aggregates → Gross Value Added by Kind of Economic Activity at constant (2015) prices - US dollars

Note: This endpoint returned constant 2015 USD values at the time data were originally downloaded (month/year). It now returns constant 2020 USD values after UNData revisions.

License / Restrictions: Open data.

`classy.csv`

This file contains country income classifications used to divide countries into High Income and Non-High Income groups.

Citation: World Bank.

URL / Access: https://databankfiles.worldbank.org/public/ddpext_download/site-content/CLASS.xlsx

Access Date: 12/05/2025

Download Steps: Download classification CSV from World Bank link provided.

License / Restrictions: Open data.

5. Expected Outputs

The script exports separate files for paper tables, including:

- `Table\_1\_Sectoral\_exposure\_measures.xlsx`
- `Table\_2\_Main\_results\_global\_sample.xlsx`
- `Table\_3\_Main\_results\_income\_groups.xlsx`
- `Table\_4\_Placebo\_test\_energy\_input\_intensity.xlsx`
- `Table\_5\_Leave\_one\_sector\_out\_robustness.xlsx`
- `Table\_6\_Appendix\_Descriptive\_statistics\_by\_income\_group.xlsx`

6. Runtime

Expected runtime is approximately 1–5 minutes on a standard computer.

7. Verification Checklist

- ☐ All files are in the same folder
- ☐ Required packages installed
- ☐ Script runs successfully
- ☐ Output files created in `outputs/tables/`
- ☐ Results match manuscript tables

8. Contact

For replication questions, please contact the paper authors at evergaracobos@worldbank.org