

Contents

1. Overview
2. Data Availability
3. Instructions for Replicators
4. List of Exhibits
5. Requirements

Overview

This project reproduces the graph data underlying the network figures of Figures 6, 9, and 10. It consists of a number of Jupyter notebooks and input data files. The final network figures were created in Gephi using the graph data. The network layout algorithms in Gephi are probabilistic and will not look exactly like the paper. The original Gephi file used to create Figure 9 is also included in this package. Note that Figures 6 and 10 will not be produced without access to restricted data files.

In the top folder, you will find this README, the final manuscript (Five Frictions - Key labor market barriers to unlocking job growth in the green transition.pdf), the `envir_ff.yml` file specifying the python packages required, plus `main.py`, the main script. It also includes the following folders: - Code: Jupyter notebooks to generate results (will be executed by `main.py`) - Data: Original data sources from ESCO (in Data → ESCO), the Georgia LFS (in Data → Georgia), the Brazil shapefiles (in Data → Brazil shapefile), and restricted data on Brazil (in Data → Brazil restr). Note that the restricted data on Brazil is not included in the public version of this package. - Data_intermediate: Analysis output created for intermediate use during the main execution will be stored here (ESCO-specific files in Data_intermediate → ESCO) - notebook_out: Output notebooks will be created during the main execution - Out: Final output folder (also includes the final figure (figure 9) plus the original Gephi file)

Data Availability

Some data cannot be made publicly available in this package, specifically 2018 RAIS data (Relação Anual de Informações Sociais: Annual Report of Social Information) and derived data files. RAIS data can be requested from IBGE (<https://ces.ibge.gov.br/base-de-dados/metadados/mte/relacao-anual-de-informacoes-sociais-rais>). Without the restricted data, the package runs but will not produce Figures 6 and 10.

Data Sources

All required data is available in the Data folder

- The Georgia Labor Force Surveys (in Data → Georgia) were obtained via the website of the National Statistics Office of Georgia: <https://www.geostat.ge/en/modules/categories/130/labour-force-survey-databases> – Access year: 2025
- The ESCO data on occupations and skills (in Data → ESCO) were obtained via the European Union's ESCO (European Skills, Competences, and Occupations) website: <https://esco.ec.europa.eu/en> – Version used: 1.2.0 – Language: English – Access year: 2025
- The shapefile for Brazil (in Data → Brazil shapefile) was obtained from IBGE: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html?edicao=27411> – Version: Brasil - Municípios - 2017 – Access year: 2025
- `cnae_ff.csv` (in Data) was created manually by the authors to identify fossil fuel industries.

Statement about Rights

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

Instructions for Replicators

Ensure all required software and dependencies are installed as listed in the Requirements section.

The main.py script runs all analysis automatically. New users should follow these steps to run the package successfully: 1) Open Anaconda prompt and navigate to the main reproduction package directory 2) create a new environment: “conda env create -file=envir_ff.yml” 3) activate environment: “conda activate envir_ff” 4) make environment available for papermill: “jupyter kernelspec install -user -name envir_ff” 5) run main script: “python main.py”

The provided code reproduces selected tables and figures in the paper, as explained and justified below

No.	Exhibit name	Output filename	Script	Note
1	Figure 6a	Out/Figure 6a.png/svg	main.py	—not generated from code due to data restrictions— -
2	Figure 6b	Out/Figure 6b.png/svg	main.py	—not generated from code due to data restrictions— -
3	Figure 6c	Out/Figure 6c.png/svg	main.py	—not generated from code due to data restrictions— -
4	Figure 6d	Out/Figure 6d.png/svg	main.py	—not generated from code due to data restrictions— -
5	data for Figure 9a	Out/Figure 9a network.graphml	main.py	——
6	data for Figure 9b	Out/Figure 9b network.graphml	main.py	——
6	Figure 9	Out/similarity_transition_networks.gephi		cannot be automated (see instructions below)

No.	Exhibit name	Output filename	Script	Note
7	Figure 10	Out/Figure 10.pdf	main.py	—not generated from code due to data restrictions—

The `.graphml` files for Figures 9a and 9b are **not yet the final figures** as shown in the manuscript, which were created in **Gephi**.

Gephi cannot be automated, and the following steps must be performed **manually**.

Note: The layout algorithm in Gephi is *probabilistic*, so the final graph may look slightly different on each run. However, it should remain *qualitatively similar*.

Step-by-Step Instructions

1. Open Gephi

- Click: **File** → **Open**
- Navigate to the `.graphml` file.
- Press **OK** on the *import report* to load the network.

2. Color

- In the **Appearance** tab (top left):
 - Under **Nodes** and the painter's **color palette** icon, click the **Ranking** tab.
 - Choose **gender** as the ranking variable.
 - Manually adjust colors to match the manuscript colors if needed.
 - Click **Apply**.

3. Size

- In the **Appearance** tab (top left):
 - Under **Nodes** and the **node size** icon, click the **Ranking** tab.
 - Choose:
 - * **employment** → for *Figure 9a*
 - * **empl** → for *Figure 9b*
 - Set:
 - * *Figure 9a*: Min size = 20, Max size = 80
 - * *Figure 9b*: Min size = 40, Max size = 150
 - In the **Spline** function (bottom left of Appearance tab), pick the **third template** and press **Close**.
 - Click **Apply**.

4. Layout

- In the **Layout** tab (bottom left):
 - Select **ForceAtlas 2**.
 - Press **Run**, wait until the layout stabilizes, then press **Stop**.

- For *Figure 9b*:
 - Select **Noverlap**.
 - Press **Run**, wait for stability, then press **Stop**.

5. Exporting the Figures

- Go to the **Preview** tab (top left).
- In the left-side menu:
 - Check **Rescale weight**.
 - Set **Max. rescaled weight** to:
 - * 20 $\rightarrow$ *Figure 9a*
 - * 50 $\rightarrow$ *Figure 9b*
 - Uncheck **Curved**.
 - Press **Refresh**.
- To export:
 - Click **Export: SVG / PDF / PNG** (bottom left).

Result:

The figures should now look like the ones presented in the manuscript.

Software Requirements

- Anaconda (Python distribution and package manager): <https://anaconda.org/>
- Gephi (for Figure 9): <https://gephi.org/> – Version used: 0.10.1