

Instructions for: “Intergenerational scars: The impact of parental unemployment on individual health later in life”

This folder is used to replicate all results presented in the paper “*Intergenerational scars: The impact of parental unemployment on individual health later in life*” by Ubaldi and Picchio (2024).

The paper uses data from German Socio-Economic Panel (GSOEP), held by DIW Berlin. As there are many versions of the GSOEP, it is important to notice that the version data used in Ubaldi and Picchio (2024) is the 36th Core Version, EU edition. Variables may change name over time. Thus, it is strongly recommended to check documentations from each version to another to spot possible discontinuities. For more details about the 36th version, please visit the following links:

1. https://www.diw.de/en/diw_01.c.814095.en/edition/soep-core_v36eu_data_1984-2019_eu_edition.html
2. <https://companion.soep.de/Data%20Structure%20of%20SOEPCore/Editions.html>

The folder contains seven building blocks plus a regression .do file. The seven blocks are divided as follows:

1. **b1_gross_merging.do**, merges the relevant initial raw SOEP datasets;
2. **b2_cleaning.do**, provides a preliminary cleaning of the variables used in the paper;
3. **b3_creating_kids.do**, creates the kids dataset that will be merged in block (5);
4. **b4_creating_parents.do**, creates three separate datasets. The first is for mothers only, the second is for fathers only, and the third combines mothers and fathers cohabitating in the same household in a given period;
5. **b5_merge_families.do**, combines blocks (3) and (4);
6. **b6_analysis_sample.do**, creates the list of variables used in Ubaldi and Picchio (2024).

Once the dataset is set up, the analysis can be performed by running the .do file, **r_regression.do**. Other than allowing to replicate the results, **r_regression.do** is already set up to store the results reported in Ubaldi and Picchio (2024). The storing process works according to the following scheme:

- *C:/your_folder_path/*
- **tables**, it stores the results reported in Tables 1-5, A.1, B.1 and B.2
 - **hetero**, stores the results reported in Section 4.2 (Heterogeneity Analysis)
 - **ages**, stores the results reported in Table 7
 - **dauson**, stores the results reported in Table 9
 - **eastwest**, stores the results reported in Table 11
 - **motpat**, stores the results reported in Table 8
 - **reform**, stores the results reported in Table 10
 - **media**, stores the results reported in Table 6
 - **robustness**, stores the results reported in Section 4.3 (Sensitivity Analysis)
 - **1991birth**, stores the results reported in Table C.4
 - **altmatch**, stores the results reported in Table C.2
 - **cross**, stores the results reported in Table C.3
 - **genunemp**, stores the results reported in Table C.5
 - **waves**, stores the results reported in Table C.1

Start by selecting a path in your device where to locate the .do and .dta files needed for the replication. Before running **r_regression.do**, create a directory named “*tables*”. Once “*tables*” is created, create three subdirectories named “*hetero*”, “*media*” and “*robustness*”.

In the “*hetero*” subdirectory, create five more directories named “*ages*”, “*dauson*”, “*eastwest*”, “*motpat*” and “*reform*”. In the “*robustness*” subdirectory, create other five directories named “*1991birth*”, “*altmatch*”, “*cross*”, “*genunemp*” and “*waves*”. Once the folder tree is set, run **r_regression.do**.

r_regression.do will store all the results for the main analysis with the corresponding descriptive statistics in the “*tables*” directory. In Ubaldi and Picchio (2024), these results are reported in Tables 1-5 of the main text, and Tables A.1, B.1 and B.2 of the Online appendix. Further, it will store the .gph files of the density plots reported in Online appendix A (Figure A.1). **r_regression.do** will store the results for the mediating channels analysis (reported in Table 6 of the main text) inside the “*media*” subdirectory. Finally, the remaining results reported in sections 4.2 (i.e. Heterogeneity Analysis) and 4.3 (i.e. Sensitivity Analysis) will be stored inside the “*hetero*” and “*robustness*” subdirectories.

References

Ubaldi, M., Picchio, M., (2024). Intergenerational scars: The impact of parental unemployment on individual health later in life, *German Economic Review*, forthcoming.