This README file describes the data and code that can be used to replicate the results in the paper:

Brandén, Gunnar, Martin Nybom, and Kelly Vosters. *Like Mother, like Child? The Rise of Women's Intergenerational Income Persistence in Sweden and the United States*.

Swedish Administrative Data

The paper relies on restricted-access individual-level data from Sweden. Due to the sensitive nature of these data, they are confidential and protected according to 24 ch. 8§ of the Public Access to Information and Secrecy Act (2009:400). We are therefore prohibited from posting the data online.

Researchers interested in obtaining access to the register data used in the paper are required to submit a written application to gain approval from Statistics Sweden. The application must include a detailed description of the proposed project, its purpose as well as a description of the required datasets, variables, and analysis population. Questions regarding data access should be directed to Statistics Sweden directly (Statistics Sweden, Box 24300, 10451 Stockholm, Sweden; mikrodata.individ@scb.se).

Researchers can apply for access to the final working data set used in the study from the Institute for Evaluation of Labour Market and Education Policy (IFAU). IFAU is not entitled to distribute the raw data, but can provide access to the analysis data file for replication purposes. IFAU decides, after examination pursuant to 24 ch. 8§ of the Public Access to Information and Secrecy Act (2009:400), to disclose the analysis data with a reservation of secrecy in accordance with 10 ch. 14§ of the Public Access to Information and Secrecy Act. Access to the data will to some extent depend on the location of the researcher and must be handled on a case-by-case basis. Contact IFAU (IFAU, Box 513, 75120 Uppsala, Sweden; ifau@ifau.uu.se) for more details on the application.

The combined data set is held by IFAU and have been accessed through servers protected by single-use passwords and encryption. We here provide and explain the programs used to generate our samples and results; the directories should be adjusted to as needed. We will also cooperate in any efforts to replicate our results.

SWE_MASTER.do is a master do-file that will replicate all results for Sweden presented or discussed in the paper, by executing the following files.

Create analysis datasets

- prep.do: prepares the raw data for analysis
- swepanel_100usd.do: generates an income panel with incomes > 100USD
- swepanel_zeros.do: generates an income panel including zeros

Results for main IRP trends (Figures 1-4) and assortative mating patterns (Figure 5):

- fig1-4_swe.do: Creates dta file for IRP trends and actual SWE Figures 1-4
- descr_byperiod_swe.do: Creates descriptive statistics dta file for SWE part of Table 1
- fig5_figA1_swe.do– Creates Figure 5 and Figure A1

Results for the main decomposition (SWE dta files for Figures 6-7) are created using:

- SE_model_5b_sequential_parent_AM.do (maternal IRP)
- SE_model_5bF_sequential_parent_AM.do (paternal IRP)
- SE_model_5bavg_sequential_parent_AM.do (household income IRP)

Results for mother-daughter employment transmission (SWE dta files for Figure 8):

- SE_model_5b_daughters_empu.do (non-employment part)
- SE_model_5b_daughters_empxb.do (employment part)

Results for exploring causal vs. measurement error explanation (SWE dta files for Table 2):

- irps_causal_vs_me.do
- irps_causal_vs_me_incomesIV.do
- irps_causal_vs_me_eduempIV.do

Appendix A results (Figures A2, A5, A6, A7, A9)

- figsappendix_swe.do

Appendix B decomposition (SWE dta files for simple model):

- SE_model_2b_sequential_parent-only.do
- SE_model_2bF_sequential_parent-only.do

The results .dta datasets are provided in the swe_results folder.

US Panel Study of Income Dynamics

The US data for this study are from the public use data from the Panel Study of Income Dynamics (PSID); the published version of the paper uses packaged raw data downloaded on August 24, 2021. These data are produced and distributed by the Institute for Social Research at the University of Michigan, Ann Arbor, MI: https://psidonline.isr.umich.edu/.

Following the PSID's data use agreement, we have deposited the raw downloaded data and analysis sample data sets in the PSID Public Data Extract Repository maintained through the Open Inter-university Consortium for Political and Social Research (ICPSR) initiative. The data for this project can be found at https://doi.org/10.3886/E209068V1 with a title: *US analysis for "Like Mother, like Child? The Rise of Women's Intergenerational Income Persistence in Sweden and the United States."*

Instructions for replication of US results

Option 1: Replication starting from raw PSID data:

- Obtain raw PSID data (downloads.zip) and PSID-provided code (psidcode.zip) files from the ICPSR deposit at https://doi.org/10.3886/E209068V1
- Set up directories as explained and listed at the beginning of US_MASTER.do
- Run US_MASTER.do

Option 2: Replication starting from cleaned analysis sample files:

- Obtain cleaned analysis samples (analysis-sample.dta, analysis-sample-main.dta) from the ICPSR deposit at https://doi.org/10.3886/E209068V1
- Set up the required directories (code, projdata, us_results, tabfig, swe_results) as explained and listed at the beginning of the US_MASTER.do
    - Note: put analysis-sample.dta and analysis-sample-main.dta in directory assigned to *projdata* macro at start of US_MASTER.do
- Run the last two sections of US_MASTER.do. These sections are called:
    - "PSIDresults - Analysis/Results based only on US data"
    - "Tables/Figures that also use SWE results files"

Descriptions of US code files provided

US_MASTER.do is a master do-file that will replicate all results for the US presented or discussed in the paper. The directory setup is at the start of this code file; this should be modified as desired. Then the master do-file can be run and will execute the files described below.

"PSIDload" directory: Files get raw PSID data files and combine them into single Stata dataset.

- *LOADFILES.do* – Takes each of the raw PSID data files and saves them as Stata data sets
- *extract.do* – Creates individual-level dataset (with select family variables) (entire PSID)
- *m_psid.do* – Renames variables and minor cleaning of dataset created by extract.do

Note: The raw data files used here can be found in the PSID Public Data Extract Repository at https://doi.org/10.3886/E209068V1.

"PSIDprep" directory: Files clean raw PSID data and create analysis datasets for this project:

- 1_individual-clean.do – cleans PSID variables for all individuals in SEO, SRC
- accuracy-flags.do – excludes major-imputed values for relevant variables
- 2_child-parent-linked-IDs.do – assigns mother/father IDs to individuals
- 3_child-parent-merge.do – adds parents (wide) to child observations
- 4a_analysis-sample-long.do – reshapes long, so panel (individual-year obs)
- 4b_analysis-sample.do – creates measures and analysis sample datasets

Note: The analysis sample files created here (*analysis-sample.dta*, *analysis-sample-main.dta*) can be found in the PSID Public Data Extract Repository at https://doi.org/10.3886/E209068V1.

"PSIDresults" directory: Files perform analysis of PSID data, creating results datasets (.dta) and paper figures/tables that include only US results. Then, the last two files create figures/tables that combine US and Sweden results .dta datasets:

- figures_trends_us.do – Creates dta file for IRP trends and actual US Figures 1-4
- US_descr_byperiod.do – Creates descriptive statistics dta file for US part of Table 1
- figures_assortmating_us.do – Creates Figure 5

Results for the main decomposition (US dta files for Figures 6-7) are created using:

- US_model_5b_sequential_parent_AM.do (maternal IRP)
- US_model_5bF_sequential_parent_AM.do (paternal IRP)

- US_model_5bavg_sequential_parent_AM.do (household income IRP)

Results for mother-daughter employment transmission (US dta files for Figure 8):

- US_model_5b_daughters_empu.do (non-employment part)
- US_model_5b_daughters_empxb.do (employment part)

Results for exploring causal vs. measurement error explanation (US dta files for Table 2):

- US_irps_causal_vs_me.do
- US_ irps_causal_vs_me_incomesIV.do
- US_irps_causal_vs_me_eduempIV.do

Appendix A results (US tables and figures)

- appendixA_tabfig_us.do

Appendix B decomposition (US dta files for simple model):

- US_model_2b_sequential_parent-only.do
- US_model_2bF_sequential_parent-only.do

The two files below require the Swedish dta results files as inputs to create tables and figures that combine the results for the two countries:

Main text Tables 1, 2 and Figures 6, 7, 8 are produced in:

- tabfig_us_se.do

Appendix B tables and figures for the decompositions are produced in:

- appendixB_tabfig_us_se.do