Replication package for “Intergenerational Spillovers of Integration Policies: Evidence from Finland’s Integration Plans”

Overview

This document describes the replication files that reproduce the results for “Intergenerational Spillovers of Integration Policies: Evidence from Finland’s Integration Plans” by Pesola and Sarvimäki. The document also provides information on data availability, a list of all data used in the paper and a description of computational requirements and the codes used. No table or figure can be reproduced without access to the confidential data.

Data Availability and Provenance

The data is constructed using individual-level data held at Statistics Finland and publicly available data on Human Development Indices and price indices. The data sources are described below. Restricted access Statistics Finland data cannot be made publicly available. Access to these data is not restricted to any particular institution or research group but requires a license from Statistics Finland. The data can be analyzed at Statistics Finland's remote server. At the time of writing, the server is known as FIONA and is accessible from universities and research institutes located in the EU or EFTA subject to Statistics Finland's approval of the institution's data security. Details of the application process are available at: http://www.stat.fi/tup/mikroaineistot/index_en.html.

Details on Data Sources

The individual-level data for the main analysis originates from administrative registers obtained from Statistics Finland. We use both ready-made data modules provided by Statistics Finland and data constructed specifically for our study by Statistics Finland.

Data sets specific to our study:

- Data on links between children and parents: int_lapset_eta

Variable	**Description**
shnro	pseudonymized person ID
vuosi	year
ika	age
sukup	gender
svaltio	country of birth
syntyv	birth year
ashnro	mother's pseudonymized ID
ishnro	father's pseudonymized ID
asvaltio	mother's country of birth
isvaltio	father's country of birth

- Data on adult immigrants' characteristics: integration_data

Variable	Description
elig	indicator of eligibility for integration plan
agea	age at arrival
marcoA	indicator of marital status at arrival
ch18A	indicator of having children under age 18 at arrival
unempa	regional unemployment rate at arrival
fem	gender
regiona	municipality of residence at arrival
akrA	type of residence municipality (urban, semi-urban, rural)
legal2	legal status (refugee, Ingrian Finn, family member, other/unknown)

- Data on integration plans: ythaksu

Variable	Description
thspv	date of plan
kotsk	type of plan

- Data on biological and adopted parents for population: lapset_vanhemmat_eta

Variable	Description
shnro	pseudonymized person ID
isan_shnro	mother's pseudonymized ID
aidin_shnro	father's pseudonymized ID

Ready-made modules:

We use Statistics Finland's ready-made data modules on background characteristics, migration, education, income, employment and employers. Detailed descriptions of these datasets are available at:

- FOLK Basic data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872023_jua_perus24_001.xml
- FOLK Income data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872021_jua_tulo23_001korj.xml
- FOLK Migration data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19832022_jua_muutt23_001.xml (has replaced the custom-made data set used in our study)
- FOLK Degree/Qualification data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872022_jua_tutk24_001.xml
- FOLK Employment data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872021_jua_tkt24_001.xml
- FLEED combined employer-employee data https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=YA244_19882016_jua_totalhenkilot_002.xml
- EDUC Students data module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=EDUC_19992022_jua_opisk_001.xml

- EDUC Entrance to secondary education module: https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=EDUC_20142022_jua_tyhr23_001.xml (has replaced the custom-made data set used in our study)

Publicly available data

We use publicly available data on Human Development Indices, price indices and country classifications. The publicly available data originates from the following sources:

- Human Development Indices available at: https://hdr.undp.org/data-center/documentation-and-downloads
- Index for inflation adjustment available at: https://stat.fi/en/statistics/khi. We use 2019 as the base year.
- Country classification based on: https://stat.fi/en/luokitukset/valtio/valtio_27_20180101/

We include the HDI, price index and country classification data as part of the replication package (HDI_1990_2017.xlsx, eki8819.xlsx, valtioluokat.xlsx).

Computational requirements

The replication files must all be run on Statistics Finland's FIONA remote access system. The code was last run on a 64-bit operating system with 32GB of RAM using Stata/MP version 17.0. The run time will depend on the number of users on the Statistics Finland server, replicators can expect the code to run for about 12 hours. The following Stata packages were used:

- estout
- gtools
- rdob_mod2
- rdrobust

Description of programs/code

The program master.do calls the following do files in order to prepare the data and generate all the figures and tables in the paper:

Name of file	Description
master.do	Runs all the .do files listed below
settings.do	Globals and settings
data_children.do	Constructs data on children and parents for the analysis
pop_data.do	Constructs basic background data and data on natives for comparisons
degrees_earn.do	Constructs data on average earnings for different degrees
education_outcomes.do	Constructs data on education outcomes
child_labor_mkt_outcomes.do	Constructs data on children's labor market outcomes
peers.do	Constructs data on 9th grade peers
par_labor_mkt_outcomes.do	Constructs data on parents' labor market outcomes
par_colleagues.do	Constructs data on parents' colleagues
reg_data.do	Constructs data for the main analysis and robustness checks
natives_mechanisms.do	Constructs data on mechanisms for natives
results_main.do	Generates tables 1, 2 and figures 1,2,3. Note that some formatting from the Excel tables to the final latex files is necessary.
results_appendix.do	Generates tables and figures in the online appendix

Instructions to Replicators

Submit an application to Statistics Finland requesting access to the data (or equivalent) listed above. Once you have the project set up, request Statistics Finland to upload the code and data included in this replication package to the project. Adjust the macros indicating project's folder structure at the top of master.do and settings.do to fit your project. Run master.do using Stata.