

README for Replication

Status Externalities in Education and Low Birth Rates in Korea

Seongeun Kim Michèle Tertilt Minchul Yum

April 2024

1 Overview

The replication package contains three folders, `\STATA`, `\EXCEL` and `\MATLAB`. The `STATA` and `EXCEL` folders contain Stata code, Excel files and data files for the empirical part, and the `MATLAB` folder contains MATLAB code files to replicate the quantitative results in the paper. We describe details below.

2 Data Availability and Provenance Statements

2.1 Details on each Data Source

The empirical work is based on three datasets from different sources: the Korean Labor and Income Panel Study (KLIPS) from [Korea Labor Institute \(2018\)](#), the Private Education Expenditures Survey (PEES) from [Statistics Korea \(2013; 2018\)](#), and the Korean Time Use Survey (KTUS) from [Statistics Korea \(2009\)](#). The replicator has to download the data files in `dta` format and store them in the prepared subdirectories.

We used the KLIPS surveys from Wave 1 to 20, which were downloaded in August 2019 from https://www.kli.re.kr/boardDownload.es?bid=0019&list_no=133256&seq=13813. The KLIPS source files and documents are available at <https://www.kli.re.kr>. Registration is required to access the data. After registration, select "KLIPS" from the top menu, go to "DATA DOWNLOAD" and select "KLIPS 1-20th wave STATA version". We used the "non-response-corrected version" of the dataset, which is only available on the Korean website. The `STATA` folder in the replication package contains the auxiliary data files for the analysis extracted from the raw KLIPS dataset: *data_clean.dta*, *data_educost_detail_wide.dta*, *children_school.dta*, *intergeneration.dta*, and *timeuse_merge.dta*.

The PEES and KTUS datasets are not allowed to be redistributed by users. The two datasets are available from the following website: <https://mdis.kostat.go.kr/index.do>. To access the data, users must register on the website, and registration requires a valid Korean phone number for identity verification. Therefore, virtually only Korean residents can access the raw datasets. The paper uses the PEES microdata downloaded in November 2022 for the years 2009-2013 and in January 2023 for the years 2014-2018. The paper also uses the 2009 KTUS microdata downloaded in October 2022. Once the raw datasets are downloaded, *1_data_cleaning.do* in the *STATA* folder generates the following auxiliary data files: *priv_edu_mod_merged.dta* and *KTUS_2009_selected.dta*. These data files are not included in the replication package because redistribution is not allowed even for a subset of the raw microdata.

The *cpi.dta* file in the *STATA* folder contains the consumer price index for Korea from 1998 to 2018, normalized to 100 in 2012 ([Korean Statistical Information Service, 1998-2018](#)). This data is freely available at <https://kosis.kr> and was retrieved in August 2019. The replicator can select English and search for *cpi* on the website. It currently provides the CPI normalized to 100 in 2020.

Data Name	Type	File Name	Location	Provided
KLIPS	Raw data	<i>klips01h.i.dta</i> – <i>klips20h.i.dta</i> , <i>klips01p.i.dta</i> – <i>klips20p.i.dta</i> , <i>klips01a.dta</i> – <i>klips20a.dta</i>	<i>/STATA/KLIPS</i>	False
	Auxiliary data	<i>data_clean.dta</i> , <i>data_educost_detail_wide.dta</i> , <i>children_school.dta</i> , <i>intergeneration.dta</i> , <i>timeuse_merge.dta</i> .	<i>/STATA</i>	True
PEES	Raw data	<i>2009_priv_edu.dta</i> – <i>2018_priv_edu.dta</i>	<i>/STATA/PEES</i>	False
	Auxiliary data	<i>priv_edu_mod_merged.dta</i>	<i>/STATA</i>	False
KTUS	Raw data	<i>KTUS_2009.dta</i>	<i>/STATA/KTUS</i>	False
	Auxiliary data	<i>KTUS_2009_selected.dta</i>	<i>/STATA</i>	False
CPI	Raw data	<i>cpi.dta</i>	<i>/STATA</i>	True

2.2 Statement about Rights

We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript. We certify that the authors of the manuscript have permission to redistribute/publish the data contained within this replication package.

2.3 Summary of Availability

Some data cannot be made publicly available.

3 Computational Requirements

- STATA (code was last run with version 17)
 - The following packages are required: *ginidesc*, *cleanplots*, *boottest*, *estout*, *fillmissing*. The do-file “1_data_cleaning.do” installs all of these packages.
 - Running time for the empirical part was less than 10 minutes on a 4-core Intel-based laptop with Windows 11 Home 22H2 version and 8 GB memory.
 - Around 1GB of hard-disk space is required to run the code with all the raw microdata files.
- MATLAB
 - The code has been run with various versions of MATLAB. The parallel computing toolbox is required.
 - The total time expected to execute the code heavily depends on the number of CPU cores because it utilizes parallelization. It depends on the speed of each core. Estimated time given below for each task is based on MATLAB version 2023a run on a Windows 11 PC with Intel i7 12700 (12 cores, 20 threads) and 32GB ram.

4 Instructions to replicate the empirical part

The following eight do-files in the *STATA* folder generate the figures, tables, and statistics used in the empirical part.

- *master.do* is the master file and runs the following codes. The replicator needs to change the folder address in line 6.
- *1_data_cleaning.do* installs necessary Stata packages and generate the following auxiliary datasets for analysis from the raw data: *data_clean.dta*, *data_educost_detail_wide.dta*, *children_school.dta*, *intergeneration.dta*, *timeuse_merge.dta*, *priv_edu_mod_merged.dta*, and *KTUS_2009_selected.dta*.
- *2_cohort_stats.do* generates Tables A1 and A7 and Figures 1.A., 1.B., and A6, and calculates the Gini coefficient, the income elasticity of demand for children, and the childlessness rate.
- *3_educost_stats.do* generates Tables 2, A2, and A3, and Figures A1 and A2, and calculates the fraction of total life-time education spending per child in income and the income elasticity of private education expenditures.
- *4_spillover_effects.do* generates Tables 3, A4, and A5.
- *5_parallel_trends.do* generates Figures 2 and A5.
 - This code uses *priv_edu_mod_merged.dta* which is not included in the replication package. This dta file needs to be generated from the PEES microdata.
- *6_KTUS_stats.do* generates Figure A4. The replicator needs to change the folder address in line 7.
 - This code uses *KTUS_2009_selected.dta* which is not included in the replication package. This dta file needs to be generated from the KTUS microdata.
- *7_intergen_persistence.do* generates Table A6 and calculates the average intergenerational elasticity of income.
- *8_KLIPS_timeuse.do* generates Table A8 and calculates the hours of work and the childcare time.

5 Model Replication

There are two subfolders in the folder \MATLAB. All results in the paper can be reproduced by executing the batch file in each subfolder, as detailed below.

5.1 File Organization

The folder `\MATLAB\Benchmark` contains all the replication MATLAB code files to produce the results from the status externality model. A `Baseline.mat` file will be generated automatically during the step 1, which is needed for the following tasks. In addition, it contains the following files (explained in an alphabetical order):

- `GenFigureA9.m`, `GenFigureA11.m`, `GenFigureA12.m`, `GenFigureA13.m`: functions for generating figures
- `gengrid.m`: a function for generating customized grids
- `GenTableA10.m`, `GenTableA11.m`: functions for generating tables
- `ginicoeff.m`: a function for computing Gini coefficient from simulated data
- `GSS.m`: a function for conducting Golden Section Search
- `HES.m`: a function for computing household equivalence scale-adjusted consumption
- `HKprodfn.m`: a function for calculating a human capital production output
- `lini.m`: a function for computing a piecewise-linearly interpolated function value
- `ols1.m`: a function for obtaining univariate ordinary least square estimates
- `Prog_MainSS.m`, `Prog_MainSS_All.m`: program for computing steady state equilibrium
- `Prog_MainTR.m`, `Prog_MainTR_App.m`: program for computing transition equilibrium
- `qualutil10.m`: a function for calculating utility from the quality of children
- `Solve.m`: batch file to execute replication
- `tauchen.m`: a function for generating Tauchen-based discretized grid
- `Trn.m`: a function for calculating pro-natal transfers
- `wls1.m`: a function for obtaining univariate weighted least square estimates

The folder `\MATLAB\Tournament` contains all the replication MATLAB code files to produce the results from the tournament model. A `Baseline.mat` file will be generated automatically during the step 1, which is needed for the following tasks. In addition, it contains all the following files (in an alphabetical order):

- `gengrid.m`: a function for generating customized grids
- `ginicoeff.m`: a function for computing Gini coefficient from simulated data
- `GSS.m`: a function for conducting Golden Section Search
- `HES.m`: a function for computing household equivalence scale-adjusted consumption
- `HKprodfn.m`: a function for calculating a human capital production output
- `lini.m`: a function for computing a piecewise-linearly interpolated function value
- `ols1.m`: a function for obtaining univariate ordinary least square estimates
- `Prog_TourSS.m`: program for computing steady state equilibrium
- `Prog_TourTR.m`: program for computing transition equilibrium
- `qualutil10.m`: a function for calculating utility from the quality of children
- `Solve_Tour.m`: batch file to execute replication
- `tauchen.m`: a function for generating Tauchen-based discretized grid
- `Trn.m`: a function for calculating pro-natal transfers
- `wls1.m`: a function for obtaining univariate weighted least square estimates

5.2 Replication Instructions

Running the batch file `\MATLAB\Benchmark\Solve.m` reproduces the following results using the status externality model. We recommend that a replicator execute the following commands consecutively. The numbers in parentheses provide expected time required for each task (with the computing power specification mentioned above in Section 3) for reference.

1. Lines 9-38 (1.5 min): Preliminaries, Table 4 and Table 5, displayed on the command window
2. Line 41 (0.5 min): Figure 3 and Table 6, displayed on the command window
3. Line 44 (1.0 min): Table 7, displayed on the command window
4. Lines 48-49, 51-52 (10 min): Table 8, displayed on the command window
5. Lines 57-64 (24+ hours): Optimal policy, displayed on the command window
6. Lines 67-69 (10 min): Table 9 and Figure 5, displayed on the command window
7. Lines 72-77 (13 min): Table A10, displayed on the command window
8. Lines 80-83 (8 min): Table A11, displayed on the command window
9. Lines 86-88 (10 min): Figure A7, displayed on pop-up window
10. Lines 91-93 (10 min): Figure A8, displayed on pop-up window
11. Lines 96-114 (400 min): Figure A9, saved in eps format
12. Lines 117-119 (10 min): Figure A10, displayed on pop-up window
13. Lines 122-126 (20 min): Table A12, displayed on the command window, and Figure A11, saved in eps format
14. Lines 129-133 (20 min): Figure A12, saved in eps format
15. Lines 136-154 (400 min): Figure A13, saved in eps format

Next, running the batch file `\MATLAB\Tournament\SolveTour.m` reproduces the following results using the tournament model:

1. Lines 9-36 (1.5 min): Preliminaries, Table A13, displayed on the command windows, and Figure A14a, saved in eps format
2. Lines 40-41, 43-44 (10 min): Table 10, displayed on the command window
3. Lines 47-56 (24+ hours): Optimal policy, displayed on the command window
4. Lines 59-61 (10 min): Table 11, displayed on the command window
5. Line 64 (2 min): Figure A14b, saved in eps format

6 List of Tables, Figures, and Statistics

The following table summarizes the figures, tables, and statistics, along with their relevant program location and their output format. When output files are exported by do files, they are automatically saved in the folder \STATA\output.

Unlike the others, Figure 4 and Figure A3 are generated by the authors using Excel. The Excel files (Figure4.xlsx and FigureA3.xlsx) are located in the folder \EXCEL and no code is required to generate these figures.

Item	Page (manuscript)	Program	Line (code)	Output
Figure 1	7	2_cohort_stats.do	186, 199	Figure_1_A.pdf, Figure_1_B.pdf
Figure 2	12	5_parallel_trends.do	115	Figure_2.pdf
Figure 3	23	\MATLAB\Benchmark\Solve.m	41	Pop-up window
Figure 4	27	\EXCEL\Figure4.xlsx		Figure4.xlsx
Figure 5	35	\MATLAB\Benchmark\Solve.m	67-69	Pop-up window
Figure A1	A-3	3_educost_stats.do	265	Figure_A1.pdf
Figure A2	A-5	3_educost_stats.do	132	Figure_A2.pdf
Figure A3	A-7	\EXCEL\FigureA3.xlsx		Figure_A3.xlsx
Figure A4	A-9	6_KTUS_stats.do	182, 202, 222, 319	Figure_A4_A.pdf, Figure_A4_B.pdf, Figure_A4_C.pdf, Figure_A4_D.pdf
Figure A5	A-10	5_parallel_trends.do	116, 161, 206, 253	Figure_A5_A.pdf, Figure_A5_B.pdf, Figure_A5_C.pdf, Figure_A5_D.pdf
Figure A6	A-15	2_cohort_stats.do	348, 361	Figure_A6_A.pdf, Figure_A6_B.pdf
Figure A7	A-24	\MATLAB\Benchmark\Solve.m	86-88	Pop-up window
Figure A8	A-25	\MATLAB\Benchmark\Solve.m	91-93	Pop-up window

Figure A9	A-27	\MATLAB\Benchmark\Solve.m	96-114	FigureA9.eps
Figure A10	A-28	\MATLAB\Benchmark\Solve.m	117-119	Pop-up window
Figure A11	A-29	\MATLAB\Benchmark\Solve.m	122-126	FigureA11.eps
Figure A12	A-30	\MATLAB\Benchmark\Solve.m	129-133	FigureA12.eps
Figure A13	A-30	\MATLAB\Benchmark\Solve.m	136-154	FigureA13.eps
Figure A14	A-31	\MATLAB\Tournament\Solve_Tour.m	9-36, 64	FigureA14a.eps, FigureA14b.eps
Table 1	8	generated directly from the data source		
Table 2	9	3_educost_stats.do	275	Tables.xlsx
Table 3	11	4_spillover_effects.do	354, 463	Table3.rtf
Table 4	19	\MATLAB\Benchmark\Solve.m	38	Command window
Table 5	22	\MATLAB\Benchmark\Solve.m	38	Command window
Table 6	24	\MATLAB\Benchmark\Solve.m	41	Command window
Table 7	25	\MATLAB\Benchmark\Solve.m	44	Command window
Table 8	29	\MATLAB\Benchmark\Solve.m	48-49, 51-52	Command window
Table 9	34	\MATLAB\Benchmark\Solve.m	67-69	Command window
Table 10	38	\MATLAB\Tournament\Solve_Tour.m	40-41, 43-44	Command window
Table 11	39	\MATLAB\Tournament\Solve_Tour.m	59-61	Command window
Table A1	A-2	2_cohort_stats.do	89	Tables.xlsx
Table A2	A-4	3_educost_stats.do	285	Tables.xlsx
Table A3	A-4	3_educost_stats.do	295	Tables.xlsx
Table A4	A-8	4_spillover_effects.do	233	Tables.xlsx

Table A5	A-12	<i>4_spillover_effects.do</i>	548, 618	TableA5-1.rtf, TableA5-2.rtf, TableA5-4.rtf
Table A6	A-13	<i>7_intergen_persistence.do</i>	90	Tables.xlsx
Table A7	A-14	<i>2_cohort_stats.do</i>	209	Tables.xlsx
Table A8	A-16	<i>8_KLIPS_timeuse.do</i>	110	Tables.xlsx
Table A9	A-17	manually generated from the data source		
Table A10	A-22	<code>\MATLAB\Benchmark\Solve.m</code>	72-77	Command window
Table A11	A-23	<code>\MATLAB\Benchmark\Solve.m</code>	80-83	Command window
Table A12	A-29	<code>\MATLAB\Benchmark\Solve.m</code>	122-124	Command window
Table A13	A-31	<code>\MATLAB\Tournament\Solve_Tour.m</code>	36	Command window
Gini coefficient (Gini income)	19	<i>2_cohort_stats.do</i>	98	Statistics.xlsx
Income elasticity of fertility	6, 19	<i>2_cohort_stats.do</i>	175	Statistics.xlsx
Overall Childlessness rate	A-13	<i>2_cohort_stats.do</i>	105, 109	Statistics.xlsx
Fraction of education spending per child in income	1, 8, 19, A-5	<i>3_educost_stats.do</i>	119	Statistics.xlsx
Income elast. of investment	19	<i>3_educost_stats.do</i>	251	Statistics.xlsx

Income elast. of private education	A-3	<i>3_educost_stats.do</i>	223-232	Statistics.xlsx
Intergenerat. income elast.	19	<i>7_intergen_persistence.do</i>	100	Statistics.xlsx
Working hours	A-15	<i>8_KLIPS_timeuse.do</i>	33	Statistics.xlsx
Parental time per child	A-15	<i>8_KLIPS_timeuse.do</i>	58,59,61	Statistics.xlsx
Average number of children	A-15	<i>8_KLIPS_timeuse.do</i>	60	Statistics.xlsx

7 References

Korea Labor Institute. 2018. Korean Labor and Income Panel Study, 1998-2018. accessed: 12-August-2019, retrieved from https://www.kli.re.kr/boardDownload.es?bid=0019&list_no=133256&seq=13813

Korean Statistical Information Service (KOSIS). 2018. CPI (2012=100), 1998-2018. accessed: 14-August-2019, retrieved from https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1J20003&conn_path=I2&language=en

Statistics Korea. 2009. MicroData Integrated Service (MDIS): Time Use Survey. doi:10.23333/P.101052.001. accessed: 31-October-2022, retrieved from <https://mdis.kostat.go.kr/index.do>

Statistics Korea. 2013. MicroData Integrated Service (MDIS): Private Education Expenditures Survey, 2009-2013. doi:10.23333/P.920011.001. accessed: 18-November-2022, retrieved from <https://mdis.kostat.go.kr/index.do>

Statistics Korea. 2018. MicroData Integrated Service (MDIS): Private Education Expenditures Survey, 2014-2018. doi:10.23333/P.920011.001. accessed: 26-January-2023, retrieved from <https://mdis.kostat.go.kr/index.do>