Replication package for

"Rationed Fertility: Treatment Effect Heterogeneity in the Child Quantity–Quality Tradeoff"

Journal of Political Economy, 2025

Rufei Guo, Junjian Yi, Junsen Zhang, Ning Zhang

1. Overview

This replication package contains code to replicate all results of the paper and the online appendix. We have used Stata to clean data and run regressions, R to draw figures, and python to conduct structural estimations. We have checked functionality of code under Stata 17.0 SE, R 4.3.0, and python 3.11.5. We also provide the TikZ code of Figure 1 in the "./output" folder.

The main data sources are the 1982 and 1990 China population censuses. One of the authors obtained access to the 1% micro samples of the 1982 and 1990 China population censuses via IPUMS International in 2015. The IPUMS International prohibits redistribution of micro data extracts. Researchers can independently apply for the micro data via IPUMS International.

Another main data source is the fines on unauthorized births compiled by Avraham Ebenstein. The fines data are available at Avraham Ebenstein's webpage.

We also exploit a bundle of supplementary data sources. Access to the 2000, 2005, and 2010 China population censuses requires a collaboration with the National Bureau of Statistics of China. Researchers can apply for access to the micros samples of these censuses via the NBS-PKU Research Data Center. Section A3 of the online appendix provides detailed descriptions of all data sources.

2. Folder structure

Table 1. Folder structure of the replication package

Folder name	Content
raw	China population censuses and other raw datasets that will be used to derive the working datasets of this study.
data	Derived working datasets that will be used directly in the analysis.
dofiles	Stata dofiles for descriptive analysis and regressions
Rfiles	R markdown files to draw graphs.
pyfiles	python programs for structural estimation.
output	figures and tables produced by R and Stata.
Rdta	Stata data files (.dta) to be used in R
temp	Temporary data files.

3. Program files

The execution of program files should follow the ten steps below.

Steps 1-4. Data clean.

Step 1. Clean the primary data: the 1982 and 1990 China population censuses from IPUMS International.

- 1_c1982clean_ind_ipums.do
- 2_c1982clean_hh_ipums.do
- 3_c1990clean_ind_ipums.do
- 4_c1990clean_hh_ipums.do

Step 2. Clean supplementary data: the 2000, 2005, 2010 China population censuses from the National Bureau of Statistics of China, and control variables from various sources.

- 5_c200020052010clean_NBS_supplementary.do
- 6_clean_controls.do

Step 3. Construct policy exposure variables.

- 1_clean_fine_llf.Rmd

Step 4. Merge intermediate datasets, and compile the working sample.

- 7_clean_main_c8290hh.do

Steps 5-7. Descriptive analysis and regressions

Step 5. Descriptive analysis. (Figures 2, 3, A1, A2, A3; Table 2)

- 8_analysis_descriptive.do
- 2_analysis_descriptive_figures.Rmd

Step 6. Regressions to estimate QQ effects. (Figures A4, A5; Tables 3, A1, A5, A6, A7, A8)

- 9_analysis_reg_main.do
- 10_analysis_reg_supplementary.do
- 99_programs.do (called in the two dofiles above)
- 3_analysis_reg_figures.Rmd

Step 7. Misreported twins and parental responses (Tables A2, A3, A4, A9, A10)

- 11_analysis_misreportedtwins.do
- 12_analysis_CCTS.do

Steps 8-10. Structural estimation and decomposition

Step 8. Structural estimation of the baseline model (Table 4, Table A11)

- 1_est_main.ipynb
- 2_est_main_se.ipynb

Step 9. Decomposition based on the baseline model (Figure 4; Table 5)

- 3_decomp_main.ipynb
- 4_decomp_representative.ipynb
- 4_decomp_figures.Rmd

Step 10. Sensitivity checks (Table A12)

- 5_est_sensitivity.ipynb
- 6_decomp_sensitivity.ipynb