Replication Codes and Data

The Value of Private Schools: Evidence from Pakistan

Review of Economics and Statistics

Pedro Carneiro
University College London, Institute for Fiscal Studies and
Centre for Microdata Methods and Practice

Jishnu Das
Georgetown University

Hugo Reis
Banco de Portugal and Católica Lisbon School of Business & Economics

August 15, 2022

Overview

There are two main sets of analysis, which are discussed in turn below. The structural analysis is done in Matlab (first step estimation and policy experiments), while the OLS/IV type of analysis, namely the main tables and figures are performed in STATA. Both parts are necessary to fully replicate the paper's analysis. If you require additional information, please contact at hfreis@bportugal.pt.

1 OLS/IV analysis

Final paper analysis run in STATA/MP 16.1 for Windows (64-bit x86-64).

1.1 Stata do files

(a) Main_Tables_Figures.do - performs the analysis to obtain all tables and figures in the manuscript.

(b) Online_App_A_and_C.do - performs all tables in the online appendix A and C.

(c) Online_App_D.do - master do-file to perform all tables in the online appendix D.

(d) Data_girls_reg.do and Data_boys_reg.do - file to merge the different datasets for estimation purposes.

(e) Data_delta_aux_g.do and Data_delta_aux_b.do - file to import/copy school fixed effects variable from Matlab to obtain the dataset to run the OLS/IV estimation in STATA.

(f) Data_elast_aux_g.do and Data_elast_aux_b.do - file to import/copy elasticity variable from Matlab to obtain the dataset in STATA to calculate the elasticities.

(g) Data_cv_enr_aux_g.do and Data_cv_enr_aux_b.do - file to import/copy policy experiment variables (compensating variation and probability of enrollment in each option from the choice set) from Matlab to get the dataset in STATA to calculate policy experiment results.

(h) boot_girls.do and boot_boys.do - file to calculate bootstrapped standard errors.

1.2 Stata datasets

(a) data_sumstat_girls_pr.dta and data_sumstat_boys_pr.dta - files to run the summary statistics at individual and household level.

(b) data_sch_reg_girls.dta and data_sch_reg_boys.dta - files to run the summary statistics and OLS/IV analysis at school level.

(c) mauza_tehsil_id.dta - village and sub-districts id.

(d) school_costs.dta - school level dataset with the costs for each school.

(e) sch_dist_inst_girls.dta and sch_dist_inst_boys.dta - school distance instruments.

(f) ST_data_elast_girls and ST_data_elast_boys - auxiliary files to obtain elasticities.

(g) Elast_fig_aux_g and Elast_fig_aux_b - files to calculate elasticities.

(h) ST_data_cv_enr_noprivate_girls and ST_data_cv_enr_noprivate_boys - files to calculate the results of the No private school policy experiment.

(i) ST_data_cv_enr_voucher_girls and ST_data_cv_enr_voucher_boys - files to calculate the results of the Voucher policy experiment.

(j) boot_res_g.dta and boot_res_b.dta - auxiliary file to calculate bootstrapped standard errors.

Note: For the on-line appendix all datasets are listed in the beginning of the do-files.

2 Structural analysis

Final paper analysis run in MATLAB R2018a.

2.1 Matlab m-files

(a) mle_unobs_girls_mean.m and mle_unobs_boys_mean.m - runs the entire analysis to obtain the first step estimates in equation (5) by Maximum likelihood for girls and boys at the mean.

(b) objfun_unobs.m - Optimization routine with the objective function running inside the main m-file for both girls and boys.

(c) mktsh_test_sim_mktid.m - function to obtain market shares (inside objfun_unobs_main.m).

(d) mle_unobs_girls_p25(75).m and mle_unobs_boys_p25(75).m - runs the entire analysis to obtain the first step estimates in equation (5) by Maximum likelihood for girls at the percentile 25 and 75.

(e) Elasticities_Simulation_girls.m and Elasticities_Simulation_boys.m - run the simulations to obtain the inputs to calculate the elasticities and policy counterfactual results.

(f) mle_unobs_girls_grades.m and mle_unobs_boys_grades.m - runs the entire analysis to obtain the first step estimates for different grades.

(g) Individual_Preferences_girls.m and Individual_Preferences_boys.m - file to obtain the inputs to calculate the correlations regarding individual preferences.

(h) mle_unobs_girls_full.m and mle_unobs_boys_full.m - runs the entire analysis to obtain the first step estimates for the robustness analysis using the full specification.

(i) objfun_unobs_full.m - Optimization routine with the objective function running inside the main m-file for the full specification.

(j) elast_fee_dist_full_girls.m and elast_fee_dist_full_boys.m -routine to obtain the inputs to calculate the elasticities using the full specification.

(k) mle_unobs_girls_nchild.m and mle_unobs_boys_nchild.m - runs the entire analysis to obtain the first step estimates for the robustness analysis using the specification with number of children in the household.

(l) objfun_unobs_nchild.m - Optimization routine with the objective function running inside the main m-file for the specification with number of children in the household.

(m) elast_fee_dist_nchild_girls.m and elast_fee_dist_nchild_boys.m - routine to obtain the inputs to calculate the elasticities using the specification with number of children in the household.

(n) mle_unobs_girls_distsq.m and mle_unobs_boys_distsq.m - runs the entire analysis to obtain the first step estimates for the robustness analysis using the specification with quadratic distance term.

(o) objfun_unobs_distsq.m - Optimization routine with the objective function running inside the main m-file for the specification with quadratic distance term.

(p) elast_fee_dist_distsq_girls.m and elast_fee_dist_distsq_boys.m - routine to obtain the inputs to calculate the elasticities using the specification with quadratic distance term.

(q) Simulation_one_private_girls.m and Simulation_one_private_boys.m - run the simulation to obtain the inputs to calculate the counterfactual results for the one private school policy experiment.

2.2 Matlab datasets

(a) data_girls and data_boys - individual attributes and school level covariates with the choice set of schools for each individual.

(b) data_girls_gr34 and data_boys_gr34 - individual attributes and school level covariates with the choice set of schools for each individual eligible to be enrolled in grades 3 and 4.

(c) data_girls_gr5 and data_boys_gr5 - individual atrributes and school level covariates with the choice set of schools for each individual eligible to be enrolled in grade 5.

(d) one_priv_g and one_priv_b - auxiliary data to use on the one private school policy experiment.

Note: The description of the variables is available in the beginning of each m-file.

Instructions

The estimation procedure is iterative as explained in the article in detailed.

The main guidelines are the following:

A - Estimation results:

(i) Run mle_unobs_girls_mean.m and mle_unobs_boys_mean.m to obtain the school fixed effects estimates (delta_girls_mean and delta_boys_mean)

(ii) Import/copy delta_girls_mean and delta_boys_mean fom Matlab to Stata using Data_delta_aux_g.do and Data_delta_aux_b.do

(iii) Run Main_Table_Figure.do (or Online-App_A_and_C.do) to obtain OLS/IV estimation.

(iv) Use Table_3_4_WTP.xls (or Table_A9_A10_WTP.xls) to calculate WTP using the IV coefficents from step (iii)

B - Elasticities and Policy simulations: use IV coefficient and change accordingly in the m-files to calculate elasticities and counterfactual results.

(v) Run Elasticities_Simulation_girls.m and Elasticities_Simulation_boys.m to calculate the elasticities and/or the policy counterfactuals results.

(vi) Import/copy from Matlab the elasticities (e.g. elfee_mean_g), compensating variation (e.g. avgCV_nopr_g), and probabilities (e.g. Prob_nopr_g) variables to Stata using Data_elast_cv_g.do

(vii) Run Main_Tables_Figures.do (or Online-App_A_and_C.do) to obtain the final results according to the outcome you are interested in.

Note: Each additional alternative analysis (different percentiles or robustness) requires to execute all previous steps according to the specification you are interested in. For some of the robustness analysis you may have to change the code accordingly to execute step (i).