Medium- and Long-Term Educational Consequences of Alternative Conditional Cash Transfer Designs: Experimental Evidence from Colombia

Felipe Barrera-Osorio
Leigh L. Linden
Juan E. Saavedra

DATASET AND DO FILES

A. DATASET

The main dataset to run the final tables is the STATA data file "Final_Data_CCT_Long_Barrera-Osorio_et_al.dta". The data file contains 17,309 observations. Each row contains data at the individual level.

B. LIST OF VARIABLES

a. Baseline covariants.

SISBEN dataset is the source of these variables. Missing variables were filled with zeros to maximize power, and a dummy variable indicating the position of the replacement is included in all regressions. Maximum number of missing points in any variable is 60, out of 17,309 observations (0.0034 of the observations). Estimates do not change to the imputation due to, first, balance across the treatment groups at baseline and, second, low number of replacements.

s_teneviv_miss float   %9.0g Indicator if s_teneviv_fill was missing in the original dataset

s_teneviv_fill float   %9.0g Indicator of house ownership (1=rent, 2=own paying; 3=own paid; 4=other, 0=missing)

s_utilities_miss float   %9.0g Indicator if s_utilities_fill was missing in the original dataset

s_utilities_fill float   %9.0g Indicator of number of utilities in the household (from 0 to 6)

s_durables_miss float   %9.0g   Indicator if s_durables_fill was missing in the original dataset

s_durables_fill float   %9.0g   Indicator of durable goods in the household (from 0 to 4)

s_infraest_hh_miss float   %9.0g   Indicator if s_infraest_hh_fill was missing in the original dataset

s_infraest_hh_fill float   %9.0g   Infrastructure in the household (from 0 to 22)

s_age_sorteo_miss float   %9.0g   Indicator if s_age_sorteo_filll was missing in the original dataset

s_age_sorteo_fill float %9.0g   Age at the moment of lottery (from 1 to 85; 0 if missing)

s_sexo_miss   float %9.0g   Indicator if s_sexo_fill was missing in the original dataset

s_sexo_fill   float %9.0g   Gender of student (1=male)

s_yrs_miss   float %9.0g   Indicator if s_yrs_fill was missing in the original dataset

s_yrs_fill   float %9.0g   Years of education (from 0 to 16)

grade_miss   float %9.0g   Indicator if grade_fill was missing in the original dataset

grade_fill   float %9.0g   Grade at baseline (from 6 to 11; 0 if missing)

s_single_miss float %9.0g   Indicator if s_single_fill was missing in the original dataset

s_single_fill float %9.0g   Indicator if household is single headed (1=single)

s_edadhead_miss float %9.0g   Indicator if s_edadhead_fill was missing in the original dataset

s_edadhead_fill float %9.0g   Age of head of the household (from 19 to 98; 0 if missing)

s_yrshead_miss float %9.0g   Indicator if s_yrshead_fill was missing in the original dataset

s_yrshead_fill float %9.0g   Years of education of the head of household (from 0 to 22)

s_tpersona_miss float %9.0g   Indicator if s_tpersona_fill was missing in the original dataset

s_tpersona_fill float %9.0g   Number of individuals in the household (from 0 to 22)

s_num18_miss   float %9.0g   Indicator if s_num18_fill was missing in the original dataset

s_num18_fill   float %9.0g   Number of people under 18 in the household (from 0 to 12)

s_estrato_miss float %9.0g   Indicator if s_estrato_fill   was missing in the original dataset

s_estrato_fill float %9.0g   Strata of the household (0,1, 2 or 3 levels)

s_puntaje_miss float %9.0g   Indicator if s_puntaje_fill was missing in the original dataset

s_puntaje_fill float %9.0g   SISBEN score (from 1.45 to 21.99; 0 if missing)

s_ingtotal_miss float %9.0g   Indicator if s_ingtotal_fill was missing in the original dataset

s_ingtotal_fill float %9.0g   Income of the household in 1000 Col $ (from 0 to 4,000)

    b. Administrative and outcome variables.

m_enrolled   float %9.0g: Status of enrollment (1= Enrolled)

r_be_sisb_punt float   %9.0g. SISBEN score, based on data collected at registry.

school_code   double %10.0g. The original school code at the time of randomization. 9999 is missing

grade   float   %9.0g. Grade of student at randomization point

T1_treat   float   %9.0g. Status of treatment 1 (1=treatment)

T2_treat   float   %9.0g Status of treatment 2 (1=treatment)

T3_treat   float   %9.0g Status of treatment 3 (1=treatment)

control   float   %9.0g Indicator of control (1=control)

treatment   float   %9.0g Indicator of treatment (1=treatment)

suba   float   %9.0g Indicator of locality (1=Suba)

full_last_name float %9.0g Indicator if information of last name was available

any_ID   float   %9.0g Indicator if any ID was available

T1_treat_upper float %9.0g Indicator of treatment 1, upper grades.

control_upper float   %9.0g Indicator of control, upper grades

T1_treat_lower float %9.0g Indicator of treatment 1, lower grades

control_lower float   %9.0g Indicator of control, lower grades

enrolled_2006 float   %9.0g Indicator of enrollment, 2006

enrolled_2007 float   %9.0g Indicator of enrollment, 2007

enrolled_2008 float   %9.0g Indicator of enrollment, 2008

on_time   float   %9.0g Indicator of on-time enrollment, any year

reach_grade_9 float   %9.0g Indicator if the student reached grade 9

dropout   float   %9.0g Indicator of dropout (any year)

heldback   float   %9.0g Indicator of heldback (any year)

icfes   float   %9.0g Indicator if student presented ICFES exam

spadies   float   %9.0g Indicator if student is present in SPADIES dataset, 2006-2012

vocational   float   %9.0g Indicator of vocational tertiary study (if student is in SPADIES 2006-2012)

university   float   %9.0g Indicator of university study (if student is in SPADIES 2006-2012)

not_reported   float   %9.0g Indicator of any other tertiary study (if student is in SPADIES 2006-2012)

spadies_longterm float   %9.0g   Indicator if the students is in the SPADIES, 2016

on_time_longterm float   %9.0g   Indicator if student enrolled on time in tertiary education, long term (conditional on tertiary enrollment)

on_time_uncond_longterm float   %9.0g Indicator if student enrolled on time in tertiary education, long term, no condicional on tertiary enrollment

grad_cond_longterm float   %9.0g Graduation of tertiary education, conditional on tertiary enrollment

grad_uncond_longterm float   %9.0g 0g Graduation of tertiary education, no conditional on tertiary enrollment

C. DO FILES

All tables (Table 1-6) and appendix analysis (B1-C6) are run by the do file "Final_Table_Journal.do"