Overview

This readme file describes the dofiles and databases used to generate all the main tables and figures and appendix tables and figures in the paper “Self-, Peer-, and Teacher Perceptions under School Tracking”.

Dofile 1: generate_main_tables_figures.do

This dofile generates the main tables and figures using the two databases analysis_sample.dta and survey_response.dta.

Dofile 2: generate_appendix_tables_figures.do

This dofile generates the appendix tables and figures using the two databases analysis_sample.dta and survey_response.dta.

Dataset 1: analysis_sample.dta

Brief description

This database contains 110 variables and 47,080 observations—28,650 from students and 18,430 from teachers.

The database is at the level of (1) reporting students by (2) reporting teachers by (3) reported-on students.

This database is used to create all figures and tables except for Table A6 and Figures 1A,1B, 1C, and 1D.

Key matching variables:
id_child_reporting: identifier for each reporting student.
id_prof: identifier for each reporting teacher.
id_child: identifier for each reported-on student, using a different ID for the same student in different discontinuities.
id_child2: identifier for each reported-on student, using same ID for the same student in different discontinuities.
clasa_2: grade level, from 9 to 12.
id_liceu_disc: identifier for each discontinuity
informant: Identifier for reporters (students or teachers)
informants: Identifier for reporters (student reporter at the top-class, bottom-class, or self-reporter; or teachers)
above: Indicates whether reported-on student is assigned to the classroom above the threshold (or below)
better_class: Indicates whether reported-on student attend the classroom above the threshold (or below)

Other variables:

Student ranks

- Variables include ranks of students (normed ranks, binary ranks, and complete ranks).
- Variables: s_confid, s_effort, s_ability, s_perform, s_disrupt, s_confid_br, s_effort_br, s_ability_br, s_perform_br, s_disrupt_br, Bs_confid, Bs_effort, Bs_ability, Bs_perform, Bs_disrupt

Student characteristics

- Variables include information on students and their parents.
- Variables: media_std alphabet girl_reporting nation_reporting age_child2_reporting siblings_reporting mom_ed_reporting took_peer_bac girl nation age_child2 age_dad2 age_mom2 siblings old_sibs dad_ed mom_ed rosen_std bullying1c_std alphabet_reporting bac_peer_raw bac_rank

Teacher characteristics

- Variables include demographic information on teachers, their credentials, and position characteristics.
- Variables: t_female, t_ethnic_romanian, t_ethnic_hungarian, t_ethnic_other, t_age, t_marstat_married, t_marstat_divorced, t_marstat_never, t_marstat_other, t_children, t_educ_ba, t_educ_ma, t_educ_phd, t_educ_noba, t_cert_didact1, t_cert_didact2, t_cert_defin, t_cert_debut, t_titular, t_suplinitor, t_subj_bio, t_subj_chem, t_subj_eng, t_subj_phys, t_subj_phil, t_subj_geo, t_subj_ger, t_subj_info, t_subj_hist, t_subj_math, t_subj_mus, t_subj_rel, t_subj_rom, t_subj_pe, t_subj_oth, t_experience, t_groups_yes, t_groups_later, t_groups_no, t_homework_0, t_homework_15, t_homework_35, t_homework_45, t_homework_60, t_homework_more, t_target_bot, t_target_mid, t_target_top, t_target_all

Ancillary variables

- These variables are used to create some figures and tables. They include identifiers for different groups of participants and interactions between media score and the indicator for assignment to the above classroom.
- Variables: i1, i2, i3, i4, i5, i6, informants14, informants24, informants34, above_media_std, media_std2, above_media_std2, self_report_, reporter_higher, reporter_lower

Dataset 2: survey_response.dta

Brief description

This database contains 18 variables and 43,500 observations. All observations come from (surveyed and not surveyed) students.

The database is at the level of (1) reporting students by (2) reported-on students.

This database is used to create Figures 1A to 1D and Table A6.

Key matching variables:
id_child: identifier for each student included in the target sample of the study (= 4,350).
id_child_peer: identifier for each peer student id_child should report on.
IDC: Identifies observations when id_child is equal to id_child_peer.
id_discont: identifier for each discontinuity.
rsurvey: identifies students who responded to the survey (= 2,865).
media_std: media score, standardized.
above: Indicates whether the student is assigned to the classroom above the threshold (or below).
above_media_std: Interaction between above and media score.
any_response1: Identifies students who ranked at least one out of the 10 peers on disruptive behavior.
any_response2: Identifies students who ranked at least one out of the 10 peers on self-confidence.
any_response3: Identifies students who ranked at least one out of the 10 peers on academic effort.
any_response4: Identifies students who ranked at least one out of the 10 peers on innate ability.
any_response5: Identifies students who ranked at least one out of the 10 peers on expected performance.
all_response1: Identifies students who ranked all 10 peers on disruptive behavior.
all_response2: Identifies students who ranked all 10 peers on self-confidence.
all_response3: Identifies students who ranked all 10 peers on academic effort.
all_response4: Identifies students who ranked all 10 peers on innate ability.
all_response5: Identifies students who ranked all 10 peers on expected performance.