

Replication Files

The Matlab scripts and Stata .do files used in our analysis are listed below. Each file generates the information for the table or figure referred to in the filename (e.g., figure_2.m and figure_2.do generate Figure 2, xtable_C1.m generates Appendix Table C1, and so on). Files marked with a * require proprietary data to be executed (see section “Data” below).

The local path to the folder containing the replication files should be specified in scripts **current_directory.do** and **current_directory.m**.

Main Tables and Figures

figure_2.m figure_2.do	Monte Carlo Simulations: Impact of Economy Size on the Equilibrium Distribution of Cutoffs (Constrained/Truncated DA)
*figure_3.do	Fraction of Students Ranking Each of the Four Most Selective Schools in the Southern District of Paris, by Distance to School Cutoff
*figure_4.do	School Cutoffs in 2012 and 2013
*figure_5.m *figure_5.do	Goodness of Fit: Observed vs. Simulated Cutoffs
table_2.m	Monte Carlo Results (500 Students, 6 Schools, 500 Samples)
*table_3.do	High School Applicants in the Southern District of Paris: Summary Statistics
*table_4.do	High Schools in the Southern District of Paris: Summary Statistics
*table_5.m	Estimation Results under Different Sets of Identifying Assumptions
*table_6.m	Goodness-of-Fit Measures Based on Different Sets of Identifying Assumptions

Appendix Tables and Figures

xfigure_C1.m	Monte Carlo Simulations: Spatial Distribution of Students and Schools
xfigure_C2.m xfigure_C2.do	Monte Carlo Simulations: Equilibrium Distribution of School Cutoffs (6 schools, 500 students)
xfigure_C3.m xfigure_C3.do	Monte Carlo Simulations: Impact of Economy Size on the Equilibrium Distribution of Cutoffs (Constrained/Truncated DA)
xfigure_C4.m xfigure_C4.do	Monte Carlo Simulations: Impact of the Marginal Cost of Applying to Schools on Equilibrium Outcomes (500 Students, 6 Schools)
xtable_C1.m	Monte Carlo Simulations: Summary Statistics
xtable_C2.m	Monte Carlo Results: Unconstrained DA (500 Students, 6 Schools, 500 Samples)
*xtable_D1.do	Correlation between Student Priority Index and Probability of Ranking the Most Selective Schools in the Southern District of Paris
*xtable_D2.m	Goodness of Fit: Observed vs. Simulated Cutoffs
*xtable_E1.do	Assigned and Unassigned Students in the Southern District of Paris

Matlab Scripts (in subfolder m/)

Main Programs

MC_data_simulation.m	Simulates school choice DA under constrained DA and unconstrained DA with cost
MC_preference_estimation.m	Estimates preference based on simulated school choice data
Paris_preference_estimation.m	Estimates preference based on school choice data from Southern District of Paris
Goodness_of_fit.m	Computes goodness of fit statistics using estimated preferences for school choice data from Southern District of Paris

Functions

DA.m	Implements Gale-Shapley Deferred Acceptance algorithm
LL.m	Computes log-likelihood function for model estimated under weak truth-telling and stability
Moment_Conditions.m	Computes moment equalities and moment inequalities
TestMR.m	Constructs marginal confidence intervals for parameter estimates based on Test MR in Bugni, Canay and Shi (2017)
QStat.m	Returns test statistics for Test MR in Bugni, Canay and Shi (2017)
Stability_test.m	Performs test of stability vs. undominated strategies based on Test RS in Bugni, Canay and Shi (2015)
Test_eq_distribution.m	To test equality of two multivariate distributions using nearest neighbor comparison
parfor_progress.m	To monitor progress of parallel for-loops
set_path.m	To set paths to the different subfolders

Companion File (in subfolder do/)

school_cutoffs.do	.do file to produce distributions of school cutoffs
-------------------	---

Data Sources

The main dataset used in the paper (*district_sud.dta*) was constructed from several administrative datasets owned by the Paris Education Authority (Rectorat de l'Académie de Paris) or the French Ministry of Education (Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche, MENESR-DEPP). These datasets contain confidential information on students. We signed an agreement with both government agencies to use and merge these datasets, which prevents us from making the data available to third parties. Researchers interested in using these datasets should contact directly the statistical offices of the Paris Educational Authority and of the French Ministry of Education.

Rectorat de l'Académie de Paris
Service Statistique Académique
12, boulevard d'Indochine
75019 Paris
FRANCE

Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche
Direction de l'évaluation, de la prospective et de la performance (MENESR-DEPP)
61-65, rue Dutot
75231 Paris Cedex 05
FRANCE

The following administrative datasets were linked using an encrypted version of the National Student Identifier (*Identifiant National Étudiant*) to produce the final dataset:

- i) AFFELNET 2012 and 2013: comprehensive data for the assignment of students to high schools in the city of Paris for the academic years 2012-2013 and 2013-2014.
- ii) BASE ÉLÈVES ACADÉMIQUE (BEA) 2012 and 2013: comprehensive register of students enrolled in Paris' middle and high schools during the academic years 2012-2013 and 2013-2014.
- iii) OCEAN DNB 2012 and 2013: middle school exam scores (*Diplôme national du brevet*) of Paris middle school students in 2012 and 2013

Our empirical analysis is restricted to students from the Southern District of Paris (*District Sud*), which contains 11 public high schools. We focus on public middle school students who are allowed to continue their studies in the academic track of upper secondary education and whose official residence is in the Southern District. We exclude those with disabilities, those who are repeating the first year of high school, and those who were admitted to specific selective tracks offered by certain public high schools in Paris (e.g., music majors, bilingual courses, etc.), as these students are given absolute priority in the assignment over other students. This leads to the exclusion of 350 students, or 18 percent of the total, the majority of whom are grade repeaters. Our data thus include 1,590 students from 57 different public middle schools, with 96 percent of students coming from one of the district's 24 middle schools.

The sources used to construct each of the variables in the final dataset are provided in the table below, together with the definitions of the variables.

Final Dataset

Unit of observation: middle school student (1,590) x high school (11).

Number of observations: 17,490.

Variable	Description	Data source
stu_id	Student identifier (from 1 to 1,590)	[AFF13]
sch_id	High school identifier (from 1 to 11)	[AFF13]
stu_female	Whether student is female	[AFF13]
stu_age	Age of student (in years)	[AFF13]
stu_highses	Whether student is high SES	[AFF13]
stu_lowinc	Whether student receives the low-income bonus	[AFF13]
stu_dnb_pct	Student's DNB composite score, in percentiles between 0 and 1	[DNB13]
stu_dnbf_pct	Student's DNB score in French, in percentiles between 0 and 1	[DNB13]
stu_dnbm_pct	Student's DNB score in math, in percentiles between 0 and 1	[DNB13]
stu_priority_pts	Student's priority index at the high school, using the original scale in points	[AFF13]
stu_priority	Student's priority index at the high school, in percentiles between 0 and 1	[AFF13]
stu_assigned	Whether student is assigned to one of the 11 high schools	[AFF13]
sch_capacity	High school capacity	[AFF13]
sch_cut_pts	High school cutoff in 2013, using the original scale in points	[AFF13]
sch_cutoff	High school cutoff in 2013, in percentiles of student priority index between 0 and 1	[AFF13]
sch_cutoff_12	High school cutoff in 2012, in percentiles of student priority index between 0 and 1	[AFF12]
sch_dnb	Mean composite DNB score of students enrolled in high school 2012, in percentiles between 0 and 1	[BEA12] [DNB12]
sch_dnbf	Mean DNB score in French of students enrolled in high school in 2012, in percentiles between 0 and 1	[BEA12] [DNB12]
sch_dnbm	Mean DNB score in math of students enrolled in high school in 2012, in percentiles between 0 and 1	[BEA12] [DNB12]
sch_highses	Fraction of high SES students in high school in 2012	[BEA12]
sch_feasible	High school is ex post feasible to the student	[AFF13]
closest_sch	High school is the closest to the student's home	[AFF13]
collocated_sch	High school is co-located with student's middle school	[AFF13]
distance	Distance between student's home and high school (in km)	[AFF13]
sch_assignment	Whether student is assigned to this high school	[AFF13]
choice_rk	Rank of high school in student's submitted ROL (0 if not included)	[AFF13]
stu_enrolment	Student's enrollment status in 2013/14 (in assigned school / in other public school / in private school / not enrolled in Paris)	[BEA13]

Data sources:

[AFF12] AFFELNET 2012 (Source: Rectorat de l'Académie de Paris);

[AFF13] AFFELNET 2013 (Source: Rectorat de l'Académie de Paris);

[BEA12] BEA 2012 (Source: Rectorat de l'Académie de Paris);

[BEA13] BEA 2013 (Source: Rectorat de l'Académie de Paris);

[DNB12] DNB 2012 (Source: MENESR-DEPP);

[DNB13] DNB 2013 (Source: MENESR-DEPP).