More Harm than Good? Sorting Effects in a Compensatory Education Program

Supplementary material

Laurent Davezies Manon Garrouste

1 Effect on Exam Scores and Enrollment in High School

The results on the probability to pass the "*Brevet*" exam can be completed with an analysis of pupils' exam scores. The exam consists of two different types of tests. The grades obtained during the school year count for 60% of the final grade; the remaining 40% come from a written exam covering three subjects: French, Mathematics, and History and Geography. Continuous assessment grades depend on the school evaluation practice. Thus they are likely to be influenced by the fact that the school is RAR or not, especially if notation practices are not independent from the class average performance. In order to analyze the effect of the program on academic achievement (and not only on the probability to pass the "*Brevet*"), it is interesting to study pupils' scores in the written exam, since these are not affected by schools inner grading practices. Analyzing written exam scores in French and in Mathematics (see Table 1) shows that there is no significant difference between pupils exogenously enrolled at RAR schools and pupils exogenously not enrolled at RAR schools, whereas a naive OLS estimation always gives a worse performance for RAR pupils.

Table 1 – Estimation of the effect of enrollment at a RAR on *Brevet* grades

	OLS				RD					
					Second stage					
	Full sample	h=0.2	h=0.3	h=0.4	h=0.2	h=0.3	h=0.4	h=0.2	h=0.3	h=0.4
Y=Exam grade (/20)										
T^{RAR}	-2.71***	-1.82***	-1.69***	-1.72***	-0.56	-0.80	-0.34	2.63	2.11	3.10
	(0.06)	(0.20)	(0.15)	(0.14)	(0.58)	(0.51)	(0.49)	(2.90)	(1.88)	(2.26)
Nbr obs	1,048,556	7,062	11,570	17,809	4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters	13,377	807	1,202	1,637	81	135	191	807	1,202	1,637
Y=French exam grade (/20)										
T^{RAR}	-1.32***	-1.07***	-0.88***	-0.78***	-0.57	-0.78**	-0.58*	0.97	1.09	1.40
	(0.05)	(0.13)	(0.11)	(0.10)	(0.40)	(0.34)	(0.34)	(1.51)	(1.08)	(1.25)
Nbr obs	1,046,039	7,062	11,570	17,809	4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters	13,376	807	1,202	1,637	81	135	191	807	1,202	1,637
Y=Math exam grade (/20)										
T^{RAR}	-1.71***	-1.31***	-1.21***	-1.05***	-0.49	-0.47	-0.80*	1.53	1.34	0.80
	(0.06)	(0.16)	(0.13)	(0.12)	(0.55)	(0.44)	(0.44)	(2.07)	(1.39)	(1.36)
Nbr obs	1,045,866	7,062	11,570	17,809	4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters	13,376	807	1,202	1,637	81	135	191	807	1,202	1,637
Y=Brevet average (/20)										
T^{RAR}	-1.88***	-1.30***	-1.13***	-1.15***	-0.10	-0.42	-0.21	1.92	1.49	2.32
	(0.04)	(0.14)	(0.11)	(0.10)	(0.38)	(0.33)	(0.36)	(2.05)	(1.31)	(1.67)
Nbr obs	1,048,859	7,062	11,570	17,809	4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters	13,378	807	1,202	1,637	81	135	191	807	1,202	1,637
Y=CC average (/20)										
T^{RAR}	-1.36***	-0.96***	-0.78***	-0.78***	0.18	-0.20	-0.14	1.41	1.05	1.78
	(0.04)	(0.13)	(0.10)	(0.09)	(0.36)	(0.31)	(0.34)	(1.53)	(0.98)	(1.31)
Nbr obs	1,048,870	7,062	11,570	17,809	4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters	13,378	807	1,202	1,637	81	135	191	807	1,202	1,637
					First stage					
Enrolling school above cutoff					0.63***	0.70***	0.67***			
					(0.18)	(0.15)	(0.12)			
Nearest school above cutoff								0.25**	0.31***	0.24***
								(0.12)	(0.11)	(0.09)
F-stat					53	38	67	16	10	11
Nbr obs					4,831	8,528	12,559	7,062	11,570	17,809
Nbr clusters					81	135	191	807	1,202	1,637

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold. On average, pupils enrolled at a RAR school in 6th grade (and whose closest junior high school is near the eligibility frontier) get a 1.69 to 1.82 point lower mean exam grade at the *Brevet* than non RAR pupils. These differences are significant at the 1% level. Pupils enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is above the eligibility frontier, have a 2.11 to 3.10 point higher mean exam grade at the *Brevet* than pupils exogenously not enrolled at a RAR. These differences are not significant.

Developing pupils' educational and professional aspirations was part of the objectives of the RAR program. The idea is to prevent pupils from schooling choices that would be driven by a lack of information. Thus it seems appropriate to not only study the treatment effects on pupils' academic performances, as measured by the "*Brevet*" scores, but also on their situations at the end of junior high school. At the end of 9th grade, do RAR pupils have a similar or better situation than other pupils?

On average, pupils who attended a RAR school in 6th grade are more often enrolled in a vocational upper secondary education track, and less often in a general upper secondary education track than other pupils (by 12 to 19 percentage points). However, this is mainly due to selection and sorting into and out of RAR, since RD estimates do not show any significant treatment effect on pupils benefiting exogenously from the program (Table 2).

Table 2 – Estimation of the effect of enrollment at a RAR on the situation 5 years later

	OLS				RD					
					Second stage					
	Full sample	h=0.2	h=0.3	h=0.4	h=0.2	h=0.3	h=0.4	h=0.2	h=0.3	h=0.4
Y=Gen high school										
T^{RAR}	-0.19***	-0.16***	-0.12***	-0.12***	0.11	0.04	-0.02	0.47	0.26	0.19
	(0.01)	(0.02)	(0.02)	(0.01)	(0.10)	(0.06)	(0.05)	(0.39)	(0.19)	(0.17)
Nbr obs	1,021,691	7,137	11,729	17,983	4,989	8,807	12,908	7,137	11,729	17,983
Nbr clusters	13,390	815	1,213	1,662	81	135	191	815	1,213	1,662
Y=Pro high school										
T^{RAR}	0.19***	0.16***	0.13***	0.12***	0.01	0.06	0.10**	-0.09	-0.08	-0.03
	(0.01)	(0.02)	(0.02)	(0.01)	(0.07)	(0.05)	(0.05)	(0.19)	(0.12)	(0.12)
Nbr obs	1,021,691	7,137	11,729	17,983	4,989	8,807	12,908	7,137	11,729	17,983
Nbr clusters	13,390	815	1,213	1,662	81	135	191	815	1,213	1,662
Y=Late (junior high)										
T^{RAR}	0.01***	0.01	-0.01	0.00	-0.13	-0.10*	-0.08	-0.33	-0.15	-0.15
	(0.00)	(0.02)	(0.02)	(0.01)	(0.08)	(0.06)	(0.05)	(0.23)	(0.11)	(0.11)
Nbr obs	1,021,691	7,137	11,729	17,983	4,989	8,807	12,908	7,137	11,729	17,983
Nbr clusters	13,390	815	1,213	1,662	81	135	191	815	1,213	1,662
					First stage					
Enrolling school above cutoff					0.63***	0.70***	0.67***			
					(0.18)	(0.15)	(0.13)			
Nearest school above cutoff								0.25**	0.31***	0.25***
								(0.12)	(0.11)	(0.09)
F-stat					52	39	70	17	10	11
Nbr obs					5,342	9,373	13,716	7,594	12,465	19,101
Nbr clusters					81	135	191	845	1,255	1,711

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold. On average, pupils enrolled at a RAR school in 6th grade (and whose closest junior high school is near the eligibility frontier) have a 12 to 16 percentage point lower probability than non RAR pupils to be enrolled in a general high school 5 years later. These differences are significant at the 1% level. Pupils enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is above the eligibility frontier, have a 19 to 47 higher probability than pupils exogenously not enrolled at a RAR to be enrolled in a general high school 5 years later. These differences are not significant.

2 Heterogeneous Effects on Academic Achievement

Table 3 – Estimation of heterogeneous effects of enrollment at a RAR on passing the *Brevet*

Y=Pass Brevet			
		RDD	
	h=0.2	h=0.3	h=0.4
Low SES (X=1) vs. High SES (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	0.56	0.60	0.67
	(0.67)	(0.56)	(0.61)
$T^{RAR}\times$*(X=1)*^{b}	0.14	0.05	0.08
	(0.10)	(0.06)	(0.08)
Test $a = b$ (pvalue)	0.490	0.313	0.313
Nbr obs	6,920	11,303	17,403
Nbr clusters	790	1,176	1,610

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. High SES pupils enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 56 to 67 percentage point higher probability to pass the *Brevet* than high SES pupils exogenously not enrolled at a RAR. This difference is not significantly different from zero. Low SES pupils enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 5 to 14 percentage point higher probability to pass the *Brevet* than low SES pupils exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Table 4 – Estimation of heterogeneous effects of enrollment at a RAR on French grade

Y=French exam grade (/20)			
		RDD	
	h=0.2	h=0.3	h=0.4
Cohort 2007 (X=1) vs. Cohort 2006 (X=0)			
$T^{RAR}\times(X=0)^a$	1.68	2.05	2.22
	(2.52)	(1.88)	(2.05)
$T^{RAR}\times(X=1)^b$	0.51	0.34	0.67
	(1.88)	(1.27)	(1.53)
Test $a = b$ (pvalue)	0.711	0.452	0.544
Nbr obs	7,068	11,583	17,830
Nbr clusters	808	1,203	1,639
Boys (X=1) vs. Girls (X=0)			
$T^{RAR}\times(X=0)^a$	1.35	1.32	2.09
	(2.13)	(1.35)	(1.56)
$T^{RAR}\times(X=1)^b$	0.60	0.82	0.90
	(1.41)	(1.07)	(1.26)
Test $a = b$ (pvalue)	0.645	0.623	0.298
Nbr obs	7,068	11,583	17,830
Nbr clusters	808	1,203	1,639
Scholarship (X=1) vs. No scholarship (X=0)			
$T^{RAR}\times(X=0)^a$	2.95	4.05	6.21
	(3.41)	(3.91)	(5.50)
$T^{RAR}\times(X=1)^b$	-0.15	0.24	0.08
	(0.72)	(0.52)	(0.55)
Test $a = b$ (pvalue)	0.358	0.321	0.258
Nbr obs	7,068	11,583	17,830
Nbr clusters	808	1,203	1,639
Low SES (X=1) vs. High SES (X=0)			
$T^{RAR}\times(X=0)^a$	3.41	4.28	4.95
	(4.99)	(4.62)	(5.01)
$T^{RAR}\times(X=1)^b$	0.32	0.31	0.32
	(0.62)	(0.46)	(0.54)
Test $a = b$ (pvalue)	0.520	0.377	0.340
Nbr obs	6,839	11,180	17,212
Nbr clusters	786	1,170	1,602

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. Girls enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, get a 1.3 to 2.1 point (over 20) higher grade in French than girls exogenously not enrolled at a RAR. This difference is not significantly different from zero. Boys enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, get a 0.6 to 0.9 point (over 20) lower grade in French than boys exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Table 5 – Estimation of heterogeneous effects of enrollment at a RAR on Maths grade

Y=Math exam grade (/20)			
	RDD		
	h=0.2	h=0.3	h=0.4
Cohort 2007 (X=1) vs. Cohort 2006 (X=0)			
T^{RAR}×*(X=0)*^{a}	3.24	2.74	1.73
	(3.76)	(2.45)	(2.23)
T^{RAR}×*(X=1)*^{b}	0.28	0.23	-0.11
	(2.36)	(1.59)	(1.63)
Test $a = b$ (pvalue)	0.505	0.392	0.507
Nbr obs	7,068	11,577	17,822
Nbr clusters	807	1,203	1,639
Boys (X=1) vs. Girls (X=0)			
T^{RAR}×*(X=0)*^{a}	1.81	1.34	1.19
	(2.91)	(1.70)	(1.61)
T^{RAR}×*(X=1)*^{b}	1.30	1.31	0.39
	(1.90)	(1.41)	(1.49)
Test $a = b$ (pvalue)	0.825	0.977	0.578
Nbr obs	7,068	11,577	17,822
Nbr clusters	807	1,203	1,639
Scholarship (X=1) vs. No scholarship (X=0)			
T^{RAR}×*(X=0)*^{a}	4.58	5.22	5.36
	(4.90)	(5.03)	(5.45)
T^{RAR}×*(X=1)*^{b}	-0.21	0.18	-0.52
	(0.93)	(0.67)	(0.73)
Test $a = b$ (pvalue)	0.331	0.315	0.283
Nbr obs	7,068	11,577	17,822
Nbr clusters	807	1,203	1,639
Low SES (X=1) vs. High SES (X=0)			
T^{RAR}×*(X=0)*^{a}	7.03	7.46	5.93
	(8.37)	(7.12)	(6.16)
T^{RAR}×*(X=1)*^{b}	-0.03	-0.05	-0.69
	(0.84)	(0.59)	(0.73)
Test $a = b$ (pvalue)	0.388	0.283	0.279
Nbr obs	6,839	11,176	17,205
Nbr clusters	785	1,170	1,602

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. Girls enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, get a 1.2 to 1.8 point (over 20) higher grade in Mathematics than girls exogenously not enrolled at a RAR. This difference is not significantly different from zero. Boys enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, get a 0.4 to 1.3 point (over 20) higher grade in Mathematics than boys exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Table 6 – Estimation of heterogeneous effects of enrollment at a RAR on enrollment in general high school

Y=Gen high school			
	RDD		
	h=0.2	h=0.3	h=0.4
Cohort 2007 (X=1) vs. Cohort 2006 (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	0.76	0.52	0.42
	(0.75)	(0.38)	(0.34)
$T^{RAR}\times$*(X=1)*^{b}	0.24	0.05	-0.03
	(0.41)	(0.19)	(0.18)
Test $a = b$ (pvalue)	0.542	0.276	0.237
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Boys (X=1) vs. Girls (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	0.61	0.27	0.19
	(0.61)	(0.24)	(0.19)
$T^{RAR}\times$*(X=1)*^{b}	0.36	0.24	0.20
	(0.30)	(0.18)	(0.19)
Test $a = b$ (pvalue)	0.550	0.860	0.928
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Scholarship (X=1) vs. No scholarship (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	0.74	0.54	0.55
	(0.67)	(0.51)	(0.56)
$T^{RAR}\times$*(X=1)*^{b}	0.23	0.15	0.09
	(0.17)	(0.10)	(0.10)
Test $a = b$ (pvalue)	0.439	0.435	0.405
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Low SES (X=1) vs. High SES (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	1.16	0.74	0.88
	(1.45)	(0.79)	(0.88)
$T^{RAR}\times$*(X=1)*^{b}	0.29	0.14	0.00
	(0.18)	(0.09)	(0.09)
Test $a = b$ (pvalue)	0.519	0.425	0.311
Nbr obs	6,898	11,314	17,347
Nbr clusters	793	1,182	1,623

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. Girls enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 19 to 60 percentage point higher probability to be enrolled in a general high school 5 years later than girls exogenously not enrolled at a RAR. This difference is not significantly different from zero. Boys enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 20 to 36 percentage point higher probability to be enrolled in a general high school 5 years later than boys exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Table 7 – Estimation of heterogeneous effects of enrollment at a RAR on enrollment in vocational high school

Y=Pro high school			
	RDD		
	h=0.2	h=0.3	h=0.4
Cohort 2007 (X=1) vs. Cohort 2006 (X=0)			
$T^{RAR}\times(X=0)^a$	-0.20	-0.27	-0.18
	(0.32)	(0.25)	(0.22)
$T^{RAR}\times(X=1)^b$	-0.01	0.08	0.10
	(0.23)	(0.13)	(0.14)
Test $a = b$ (pvalue)	0.628	0.203	0.271
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Boys (X=1) vs. Girls (X=0)			
$T^{RAR}\times(X=0)^a$	-0.05	-0.05	0.04
	(0.23)	(0.13)	(0.12)
$T^{RAR}\times(X=1)^b$	-0.10	-0.09	-0.10
	(0.20)	(0.14)	(0.16)
Test $a = b$ (pvalue)	0.771	0.700	0.297
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Scholarship (X=1) vs. No scholarship (X=0)			
$T^{RAR}\times(X=0)^a$	-0.09	-0.20	-0.20
	(0.31)	(0.31)	(0.33)
$T^{RAR}\times(X=1)^b$	-0.02	-0.03	0.01
	(0.10)	(0.07)	(0.08)
Test $a = b$ (pvalue)	0.828	0.568	0.530
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Low SES (X=1) vs. High SES (X=0)			
$T^{RAR}\times(X=0)^a$	-0.29	-0.17	-0.21
	(0.56)	(0.35)	(0.36)
$T^{RAR}\times(X=1)^b$	-0.06	-0.06	0.02
	(0.11)	(0.07)	(0.08)
Test $a = b$ (pvalue)	0.657	0.747	0.522
Nbr obs	6,898	11,314	17,347
Nbr clusters	793	1,182	1,623

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. Girls enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 4 to 5 percentage point lower probability to be enrolled in a vocational high school 5 years later than girls exogenously not enrolled at a RAR. This difference is not significantly different from zero. Boys enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 9 to 11 percentage point lower probability to be enrolled in a vocational high school 5 years later than boys exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Table 8 – Estimation of heterogeneous effects of enrollment at a RAR on repetition

Y=Late (junior high)			
	RDD		
	h=0.2	h=0.3	h=0.4
Cohort 2007 (X=1) vs. Cohort 2006 (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	-0.45	-0.17	-0.20
	(0.42)	(0.17)	(0.19)
$T^{RAR}\times$*(X=1)*^{b}	-0.22	-0.13	-0.11
	(0.24)	(0.13)	(0.13)
Test $a = b$ (pvalue)	0.633	0.863	0.687
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Boys (X=1) vs. Girls (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	-0.52	-0.21	-0.24
	(0.43)	(0.15)	(0.15)
$T^{RAR}\times$*(X=1)*^{b}	-0.19	-0.10	-0.07
	(0.16)	(0.10)	(0.11)
Test $a = b$ (pvalue)	0.322	0.337	0.181
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Scholarship (X=1) vs. No scholarship (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	-0.50	-0.17	-0.26
	(0.39)	(0.22)	(0.29)
$T^{RAR}\times$*(X=1)*^{b}	-0.20	-0.13	-0.12
	(0.14)	(0.08)	(0.09)
Test $a = b$ (pvalue)	0.421	0.873	0.618
Nbr obs	7,137	11,729	17,983
Nbr clusters	815	1,213	1,662
Low SES (X=1) vs. High SES (X=0)			
$T^{RAR}\times$*(X=0)*^{a}	-0.86	-0.49	-0.63
	(0.97)	(0.50)	(0.59)
$T^{RAR}\times$*(X=1)*^{b}	-0.17	-0.06	-0.03
	(0.13)	(0.07)	(0.08)
Test $a = b$ (pvalue)	0.454	0.360	0.293
Nbr obs	6,898	11,314	17,347
Nbr clusters	793	1,182	1,623

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h around the threshold. Girls enrolled at a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 21 to 52 percentage point lower probability to repeat a grade during junior high school than girls exogenously not enrolled at a RAR. This difference is not significantly different from zero. Boys enrolled at a RAR school exogenously, due to the fact that their closest public junior high school is just above the thresholds, have a 7 to 19 percentage point lower probability to repeat a grade during junior high school than boys exogenously not enrolled at a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

3 Analysis on Each Assignment Variable Separately

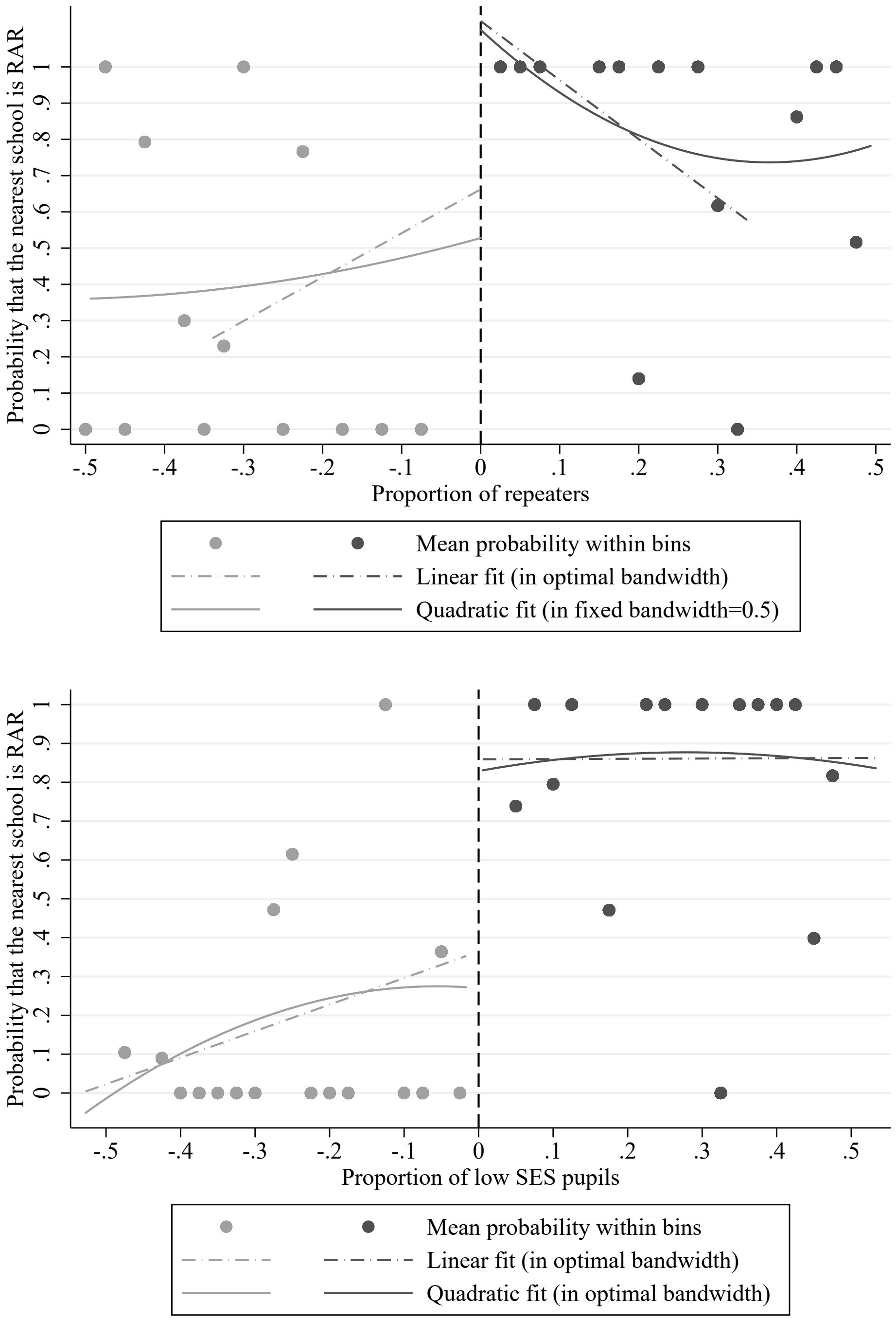

Figure 1 – Individual probability that the nearest school is RAR around each eligibility threshold

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: Each x-axis variable is centered on the respective threshold (10% for the proportion of repeaters, and 67% for the proportion of low SES), so that zero represents the eligibility threshold, and divided by its standard deviation.

Table 9 – Effect of living near a RAR school on school choice, using the 10% of repeaters assignment variable

	RD linear spline						**RD quadratic spline**					
	h=0.2	h=0.3	h=0.4	h=0.6	h=0.8	h=ob	h=0.2	h=0.3	h=0.4	h=0.6	h=0.8	h=ob
Y=Enrollment in the nearest school												
Second stage												
T^{NEAR}	-0.25**	-0.16	-0.42***	0.07	0.06	-0.09	-0.63	-0.85***	-0.37**	-0.42***	-0.07	
	(0.09)	(0.10)	(0.15)	(0.14)	(0.11)	(0.14)	(0.49)	(0.25)	(0.16)	(0.15)	(0.16)	
Mean of Y	0.47	0.46	0.43	0.44	0.44	0.42	0.47	0.46	0.43	0.44	0.44	
First stage												
$\mathbb{1}\{\frac{Z^L-0.10}{\sigma_{Z^L}} \geq 0\}$	1.00	1.50***	0.74**	0.76***	0.73***	0.94***	1.00	1.18**	1.54***	1.07***	0.98***	
	(.)	(0.19)	(0.29)	(0.21)	(0.17)	(0.26)	(.)	(0.49)	(0.42)	(0.32)	(0.30)	
F-stat	.	61	7	13	18	13	.	6	13	11	11	
Nbr obs	2,381	4,212	7,672	15,006	18,959	10,544	2,381	4,212	7,672	15,006	18,959	
Nbr clusters	28	54	80	146	188	108	28	54	80	146	188	
Y=Enrollment in another public school												
Second stage												
T^{NEAR}	0.10	0.06	-0.16	-0.17	-0.20	-0.02	-0.34**	0.68***	0.38***	0.23*	0.02	
	(0.07)	(0.09)	(0.29)	(0.15)	(0.12)	(0.13)	(0.13)	(0.21)	(0.13)	(0.13)	(0.16)	
Mean of Y	0.31	0.35	0.33	0.33	0.35	0.34	0.31	0.35	0.33	0.33	0.35	
First stage												
$\mathbb{1}\{\frac{Z^L-0.10}{\sigma_{Z^L}} \geq 0\}$	1.00	1.50***	0.74**	0.76***	0.73***	0.74***	1.00	1.18**	1.54***	1.07***	0.98***	
	(.)	(0.19)	(0.29)	(0.21)	(0.17)	(0.23)	(.)	(0.49)	(0.42)	(0.32)	(0.30)	
F-stat	.	61	7	13	18	11	.	6	13	11	11	
Nbr obs	2,381	4,212	7,672	15,006	18,959	12,829	2,381	4,212	7,672	15,006	18,959	
Nbr clusters	28	54	80	146	188	128	28	54	80	146	188	
Y=Enrollment in a private school												
Second stage												
T^{NEAR}	0.15	0.10	0.57*	0.10	0.14	0.57	0.98*	0.17	-0.01	0.19	0.05	
	(0.10)	(0.06)	(0.33)	(0.12)	(0.11)	(0.35)	(0.54)	(0.14)	(0.11)	(0.14)	(0.11)	
Mean of Y	0.23	0.19	0.24	0.23	0.21	0.23	0.23	0.19	0.24	0.23	0.21	
First stage												
$\mathbb{1}\{\frac{Z^L-0.10}{\sigma_{Z^L}} \geq 0\}$	1.00	1.50***	0.74**	0.76***	0.73***	0.74**	1.00	1.18**	1.54***	1.07***	0.98***	
	(.)	(0.19)	(0.29)	(0.21)	(0.17)	(0.31)	(.)	(0.49)	(0.42)	(0.32)	(0.30)	
F-stat	.	61	7	13	18	6	.	6	13	11	11	
Nbr obs	2,381	4,212	7,672	15,006	18,959	7,115	2,381	4,212	7,672	15,006	18,959	
Nbr clusters	28	54	80	146	188	74	28	54	80	146	188	

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at closest junior high school level. Two-stage least squares are estimated for bandwidths of size h around the 10% threshold. "ob" denotes the optimal bandwidth (Calonico et al., 2014b). Note that, due to invertibility problem, the optimal bandwidth with quadratic spline cannot be computed. This is due to a "quasi-sharp" setting close to the cutoff.

Table 10 – Effect of living near a RAR school on school choice, using the 67% of low-SES assignment variable

	RD linear spline						RD quadratic spline					
	h=0.2	h=0.3	h=0.4	h=0.6	h=0.8	h=ob	h=0.2	h=0.3	h=0.4	h=0.6	h=0.8	h=ob
Y=Enrollment in the nearest school												
						Second stage						
T^{NEAR}	-0.29	-0.44**	-0.31*	-0.06	-0.07	-0.11	0.12	-0.02	-0.33*	-0.34*	-0.16	-0.53*
	(0.20)	(0.18)	(0.18)	(0.13)	(0.11)	(0.14)	(0.19)	(0.18)	(0.18)	(0.18)	(0.17)	(0.27)
Mean of Y	0.47	0.46	0.47	0.46	0.46	0.46	0.47	0.46	0.47	0.46	0.46	0.46
						First stage						
$\mathbb{1}\{\frac{Z^F-0.67}{\sigma_{Z^F}} \geq 0\}$	0.63***	0.62***	0.46***	0.48***	0.50***	0.49***	0.86***	0.78***	0.78***	0.58***	0.51***	0.60***
	(0.21)	(0.18)	(0.15)	(0.13)	(0.11)	(0.15)	(0.30)	(0.24)	(0.22)	(0.18)	(0.16)	(0.22)
F-stat	9	12	9	14	21	12	8	11	13	11	10	8
Nbr obs	5,435	8,615	12,506	20,058	23,736	16,784	5,435	8,615	12,506	20,058	23,736	15,888
Nbr clusters	54	86	122	188	229	162	54	86	122	188	229	152
Y=Enrollment in another public school												
						Second stage						
T^{NEAR}	-0.08	0.03	-0.00	-0.05	-0.08	0.04	-0.07	-0.16	0.01	0.12	-0.01	-0.10
	(0.16)	(0.13)	(0.15)	(0.14)	(0.12)	(0.15)	(0.16)	(0.16)	(0.12)	(0.13)	(0.16)	(0.14)
Mean of Y	0.36	0.36	0.36	0.36	0.35	0.37	0.36	0.36	0.36	0.36	0.35	0.36
						First stage						
$\mathbb{1}\{\frac{Z^F-0.67}{\sigma_{Z^F}} \geq 0\}$	0.63***	0.62***	0.46***	0.48***	0.50***	0.45***	0.86***	0.78***	0.78***	0.58***	0.51***	0.73***
	(0.21)	(0.18)	(0.15)	(0.13)	(0.11)	(0.15)	(0.30)	(0.24)	(0.22)	(0.18)	(0.16)	(0.24)
F-stat	9	12	9	14	21	9	8	11	13	11	10	10
Nbr obs	5,435	8,615	12,506	20,058	23,736	14,474	5,435	8,615	12,506	20,058	23,736	10,289
Nbr clusters	54	86	122	188	229	142	54	86	122	188	229	104
Y=Enrollment in a private school												
						Second stage						
T^{NEAR}	0.37*	0.41**	0.31	0.11	0.15	0.22	-0.05	0.18	0.32**	0.22	0.17	0.31*
	(0.21)	(0.18)	(0.19)	(0.13)	(0.11)	(0.17)	(0.13)	(0.15)	(0.16)	(0.16)	(0.16)	(0.17)
Mean of Y	0.17	0.17	0.17	0.19	0.19	0.16	0.17	0.17	0.17	0.19	0.19	0.17
						First stage						
$\mathbb{1}\{\frac{Z^F-0.67}{\sigma_{Z^F}} \geq 0\}$	0.63***	0.62***	0.46***	0.48***	0.50***	0.45***	0.86***	0.78***	0.78***	0.58***	0.51***	0.80***
	(0.21)	(0.18)	(0.15)	(0.13)	(0.11)	(0.15)	(0.30)	(0.24)	(0.22)	(0.18)	(0.16)	(0.23)
F-stat	9	12	9	14	21	9	8	11	13	11	10	12
Nbr obs	5,435	8,615	12,506	20,058	23,736	14,474	5,435	8,615	12,506	20,058	23,736	9,718
Nbr clusters	54	86	122	188	229	142	54	86	122	188	229	96

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for bandwidths of size h around the 67% threshold. "ob" denotes the optimal bandwidth (Calonico et al., 2014b).

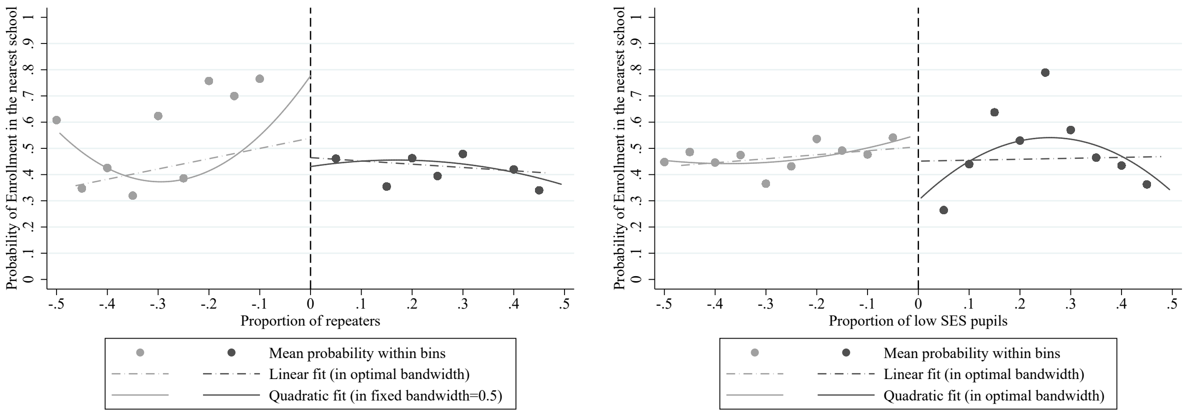

Figure 2 – Individual probability to enroll at the nearest school

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: Each x-axis variable is centered around the respective threshold (10% for the proportion of repeaters, and 67% for the proportion of low SES), so that zero represents the eligibility threshold, and divided by its standard deviation.

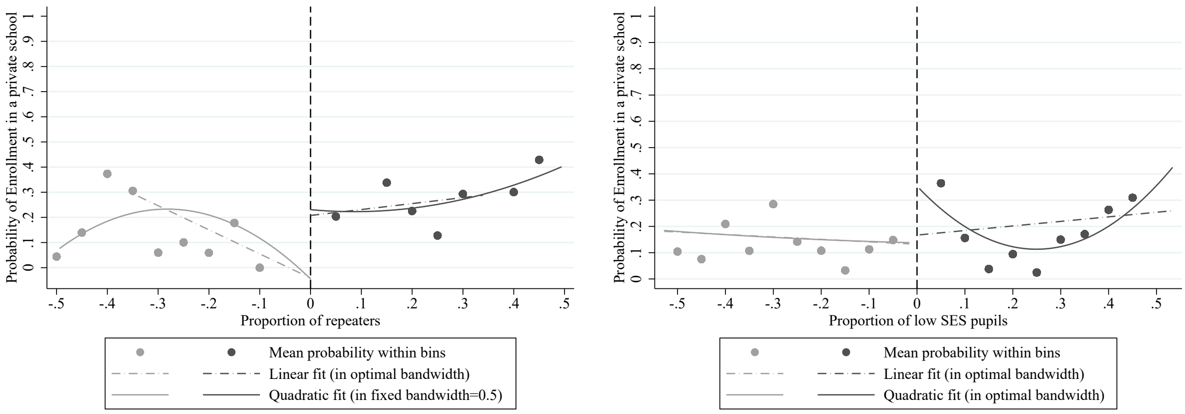

Figure 3 – Individual probability to enroll at a private school

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: Each x-axis variable is centered around the respective threshold (10% for the proportion of repeaters, and 67% for the proportion of low SES), so that zero represents the eligibility threshold, and divided by its standard deviation.

4 Robustness to the Error on the Catchment Area School

Our identification strategy relies on the use of the closest school (defined as the closest school to pupil's primary school) as a proxy for the catchment area school. Let D denote the dummy variable, which is equal to 1 if the catchment area school is the same as the "closest" school, and 0 otherwise. Let $T^*, Y(0)^*, Y(1)^*, Y^*$ and S^* respectively denote the treatment variable, potential outcomes, actual outcome, and running variable for the school of the catchment area. As soon as $D = 1$, we have that $T^* = T^{NEAR}$, $Y(0)^* = Y(0)$, $Y(1)^* = Y(1)$, $Y^* = Y$, and $S^* = S$. When $D = 0$, there is a misclassification problem. Let us assume the following:

Assumption 4.1 (Ignorable Misclassification)

1. *No almost-sure error: for any value of s in a neighborhood of 0,*

$$\mathbb{P}(D = 1|S = s) > 0,$$

$$\mathbb{P}(D = 1|S^* = s) > 0.$$

2. *Exogenous misclassification: for any value of s in a neighborhood of 0,*

$$T, Y(0), Y(1) \perp\!\!\!\perp D|S = s,$$
$$T^*, Y(0)^*, Y(1)^* \perp\!\!\!\perp D|S^* = s.$$

The first assumption ensures that, with positive probability, the closest school corresponds to the school of the catchment area. The second one states that the probability of misclassification is not correlated with treatment and potential outcomes (conditional on the running variable).

Proposition 4.2 (Robustness) *If Assumption 4.1 holds with the usual assumption of the fuzzy RD design for $T^*, Y(0)^*, Y(1)^*, S^*$ (Hahn, Todd, and van der Klaauw 2001), then both the identification and estimation of the LATE are robust to the fact that variables Y, T^{NEAR} and S are used instead of Y^*, T^* and S^*.*

Proof : For any value s in a neighborhood of the frontier, Assumption 4.1 ensures that:

$$\mathbb{E}(Y^*|S^* = s) = \mathbb{E}(Y^*|S^* = s, D = 1),$$

and

$$\mathbb{E}(Y|S = s) = \mathbb{E}(Y|S = s, D = 1).$$

By definition of D, we have $Y = Y^*$ and $S = S^*$ if $D = 1$, then:

$$\mathbb{E}(Y^*|S^* = s) = \mathbb{E}(Y|S = s).$$

A similar reasoning ensures that $\mathbb{E}(T^*|S^* = s) = \mathbb{E}(T^{NEAR}|S = s)$. Because the usual assumption of fuzzy RD design holds for T^*, Y^*, S^*, we know that $\lim_{c\downarrow 0} \frac{\mathbb{E}(Y^*|S^*\in[0;c]) - \mathbb{E}(Y^*|S^*\in[-c;0])}{\mathbb{E}(T^*|S^*\in[0;c]) - \mathbb{E}(T^*|S^*\in[-c;0])}$ converges to the LATE. So this is also the case for $\lim_{c\downarrow 0} \frac{\mathbb{E}(Y|S\in[0;c]) - \mathbb{E}(Y|S\in[-c;0])}{\mathbb{E}(T^{NEAR}|S\in[0;c]) - \mathbb{E}(T^{NEAR}|S\in[-c;0])}$.

5 Heterogeneous Effects with Respect to Local Private School Supply

Because most of parental strategies to avoid treated schools seem to be driven by pupils enrolling in the private sector, the effect is expected to differ according to the local supply of private schools. To test for this, let us define the distance to the nearest private junior high school as the smallest distance to each pupil's primary school. To test whether the effect depends on the alternative supply provided by the private sector, this measure is interacted with the treatment dummy in the model. Table 11 again shows that living near a RAR junior high school due to the eligibility thresholds significantly increases the probability to attend a private school in 6th grade, by 24 to 39 percentage points. But every additional kilometer to the neatest private school decreases this probability by 9 to 17 percentage points. The interaction term is only significant for the intermediate bandwidth however.

To further investigate this effect, let us check whether it varies with respect to individual social characteristics. Table 12 presents the results for high SES and low SES pupils. We find that the effect on enrolling at a private school of living near a RAR school decreases with distance to the nearest private school. For the sub-population of high SES pupils, every additional kilometer to the nearest private school in the neighborhood significantly decreases the probability to enroll at a private school by 11 to 22 percentage point. The interaction term is not significant for low SES pupils.

Table 11 – Estimation of the effect of living near a RAR on school choice according to distance to private school

	RD		
	h=0.2	h=0.3	h=0.4
	Second stage		
Y=Enrollment in a private school			
T^{NEAR}	0.24	0.26***	0.39**
	(0.16)	(0.09)	(0.17)
$T^{NEAR} \times$ *dist pri*	-0.17	-0.09**	-0.17
	(0.23)	(0.04)	(0.11)
Nbr obs	7,594	12,465	18,933
Nbr clusters	80	134	188

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold. Pupils living near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the eligibility frontier (h=0.3), have a 26 percentage point higher probability to enroll at a private school than pupils whose closest public junior high school is not a RAR. Every additional kilometer to the neatest private school significantly decreases this probability by 9 percentage points.

Table 12 – Estimation of heterogenous effects of living near a RAR on school choice according to distance to private school

Y=Enrollment in	a private school		
	RD		
	h=0.2	h=0.3	h=0.4
Low SES (X=1) vs. High SES (X=0)			
$T^{NEAR}\times$(X=0)^{a}	0.372**	0.390***	0.464***
	(0.186)	(0.113)	(0.170)
$T^{NEAR}\times$(X=1)^{b}	0.086	0.136**	0.314*
	(0.089)	(0.063)	(0.178)
$T^{NEAR}\times$(X=0)$\times$ *dist pri*^{c}	-0.223	-0.112**	-0.178**
	(0.260)	(0.044)	(0.082)
$T^{NEAR}\times$(X=1)$\times$ *dist pri*^{d}	-0.078	-0.063	-0.148
	(0.135)	(0.048)	(0.148)
Test $a = b$ and $c = d$ (pvalue)	0.071	0.029	0.445
Nbr obs	7,342	12,017	18,248
Nbr clusters	80	134	188

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Notes: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold. High SES pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the eligibility frontier (h=0.3), have a 39 percentage point higher probability to enroll at a private junior high school than high SES pupils whose nearest junior high school is not a RAR exogenously, and every additional kilometer to the nearest private school significantly decreases this effect by 11 percentage points. Low SES pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the eligibility frontier (h=0.3), have a 14 percentage point higher probability to enroll at a private junior high school than low SES pupils whose nearest junior high school is not a RAR exogenously, and every additional kilometer to the nearest private school decreases this effect by 6.3 percentage point, although not significantly. The difference between both pairs of estimates is jointly significantly different from zero.