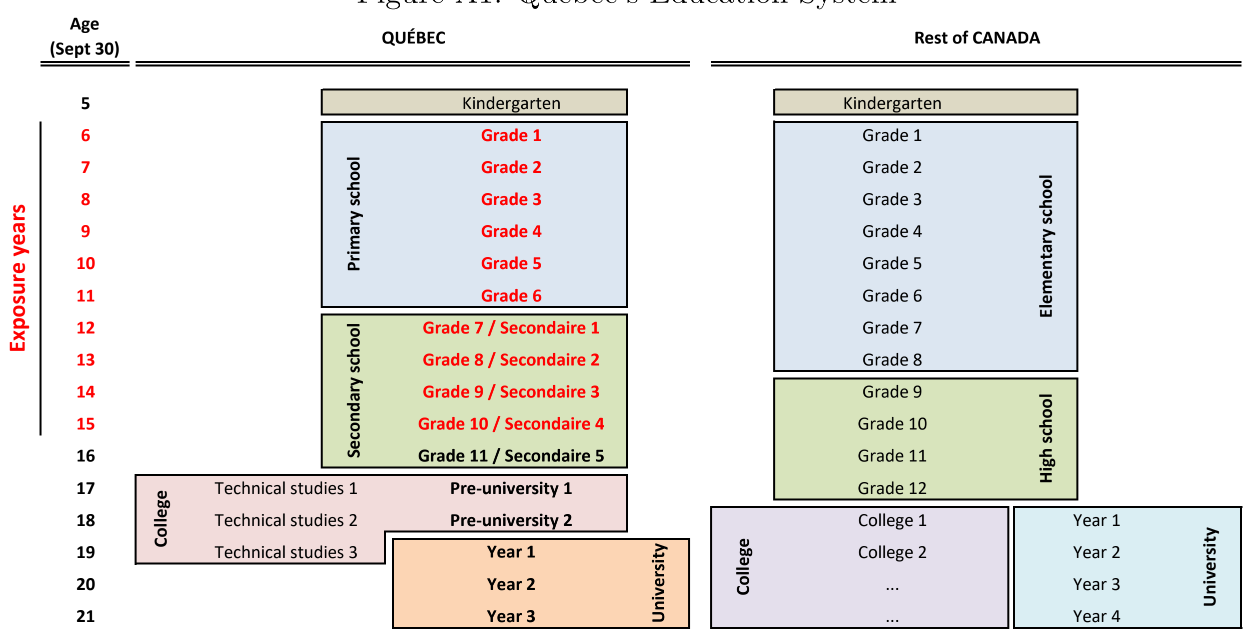
Online Appendix
"Long-term Contextual Effects in Education: Schools and Neighborhoods"

Jean-William Laliberté

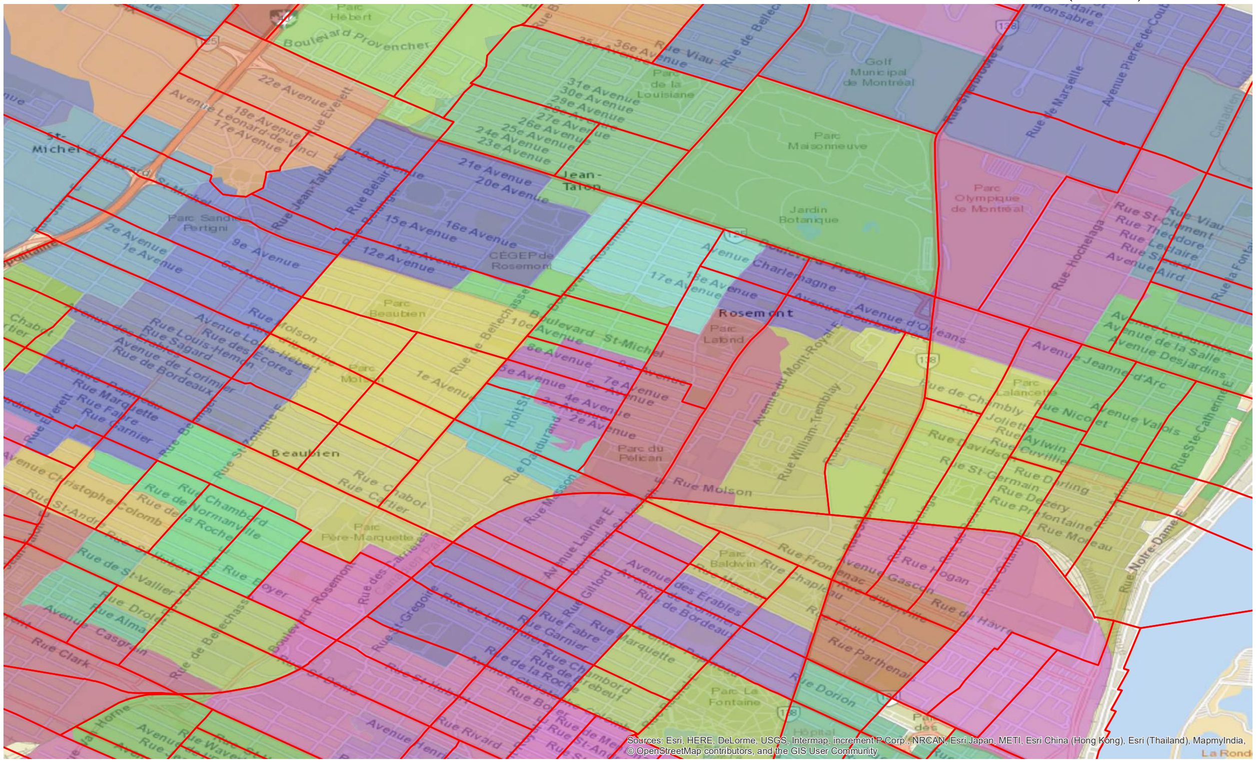
A Appendix Tables and Figures

Figure A1: Quebec's Education System



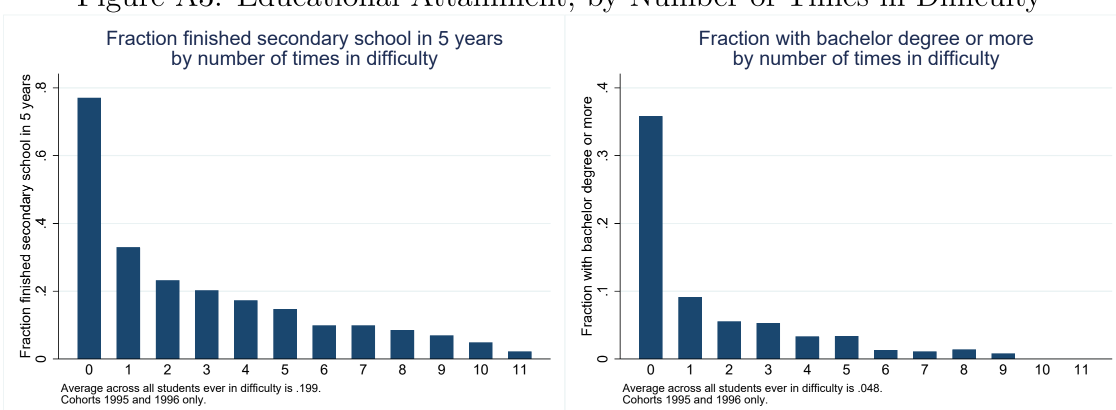
Notes: The figure contrasts Quebec's education system with a typical education system in the rest of Canada.

Figure A2: Catchment Areas and Census Tracts in Eastern Montreal (2001)



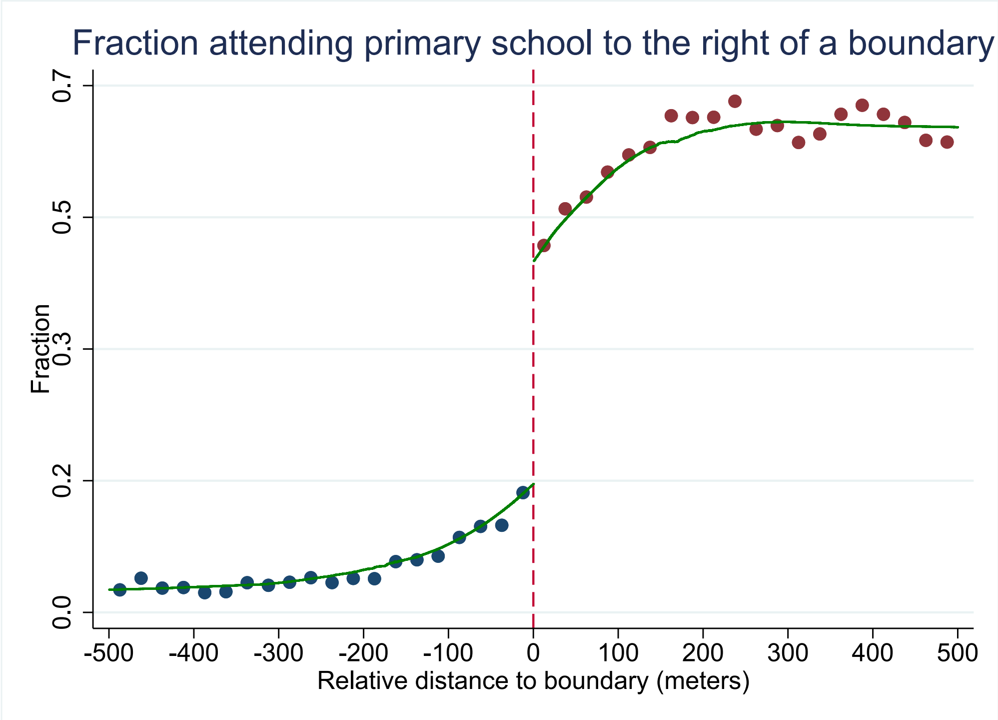
Notes: This map shows French primary school boundaries and census tract boundaries as of 2001 in Eastern Montreal. Colored areas indicate French primary school catchment areas as of 2001. Red lines denote census tracts boundaries.

Figure A3: Educational Attainment, by Number of Times in Difficulty



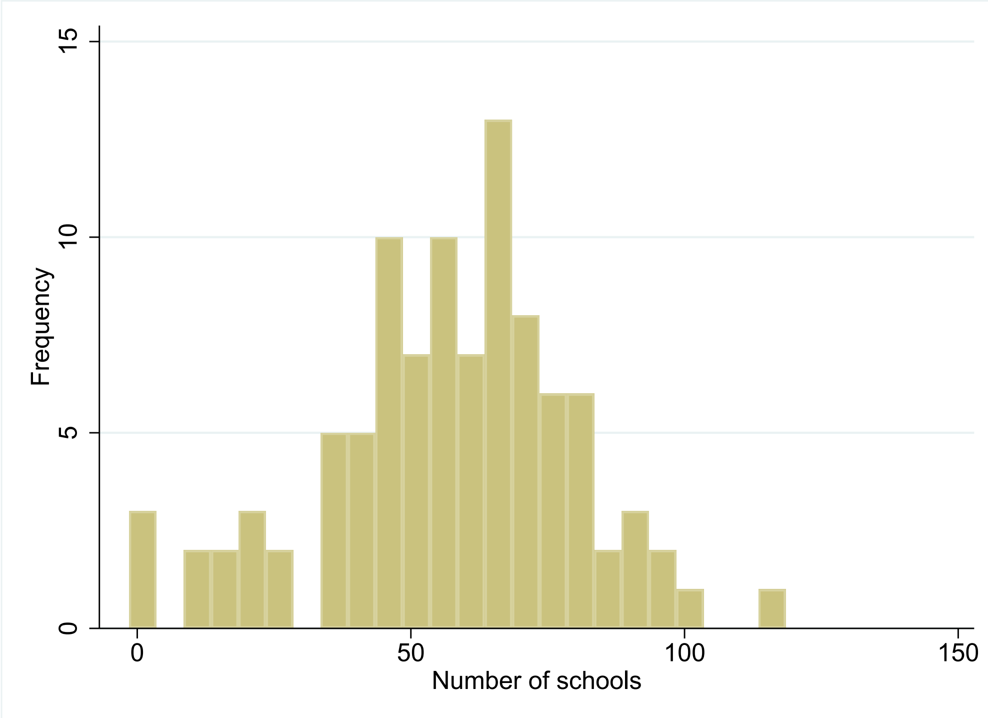
Notes: This figure shows the fraction of students who attain certain levels of education (finished secondary school on time in the left panel, and obtaining a bachelor degree on the right panel) as a function of the number of times they are identified as being “in difficulty” before the age of 16. The sample is restricted to students from the 1995 and 1996 cohorts. Too few students of the later cohorts have completed a bachelor degree by 2014-2015 to analyze this outcome for these students.

Figure A4: Discontinuity in School Attendance At Boundaries



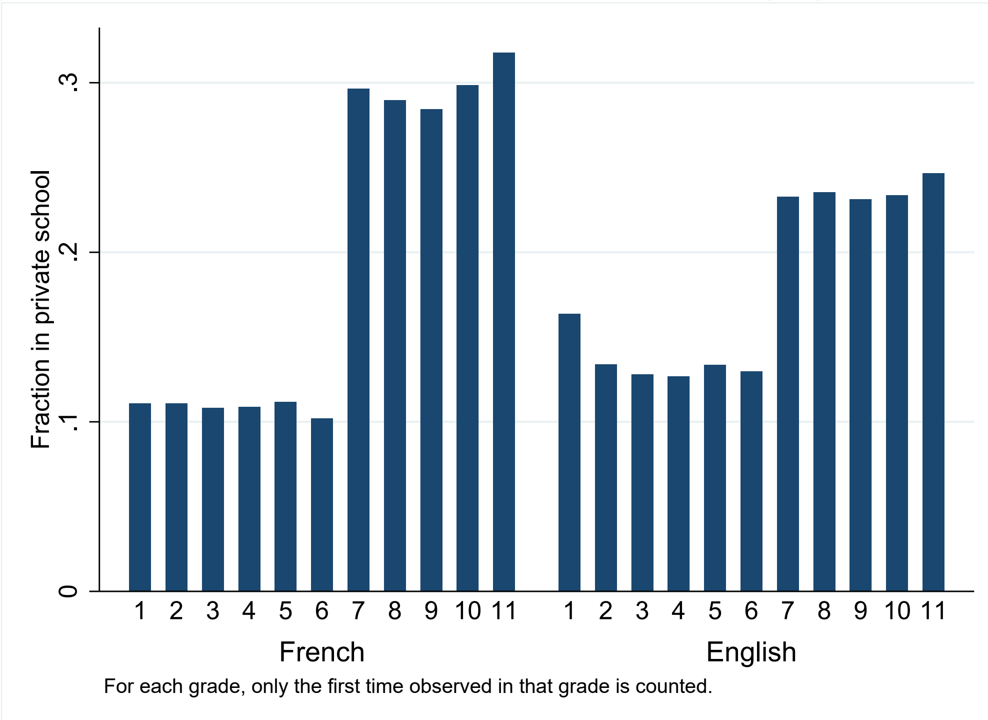
Notes: For each French primary school boundary, one of the two adjacent neighborhood schools is randomly chosen and labeled as being "to the right of a boundary". The figure shows the fraction of students enrolled in that randomly chosen school, by distance to the boundary. Students at positive distance are those assigned to the randomly chosen default school. Students at negative distances are assigned to a different default school (i.e. one other than the one to the right). For visual clarity, students living further than 500 meters away from their nearest boundary are excluded. Each dot indicates the share of students attending the randomly chosen neighborhood school within 25-meter bins. Attendance is recorded at baseline (first enrollment in grade 1). The sample is restricted to students in French schools.

Figure A5: Distribution of School Choice Across FSAs



Notes: This histogram shows the distribution of FSAs by number of different primary schools attended by the students who reside in them. Statistics calculated over main analytical sample of 92,764 students who completed their compulsory schooling in Montreal. School attendance is measured at baseline (i.e. first enrollment in grade 1).

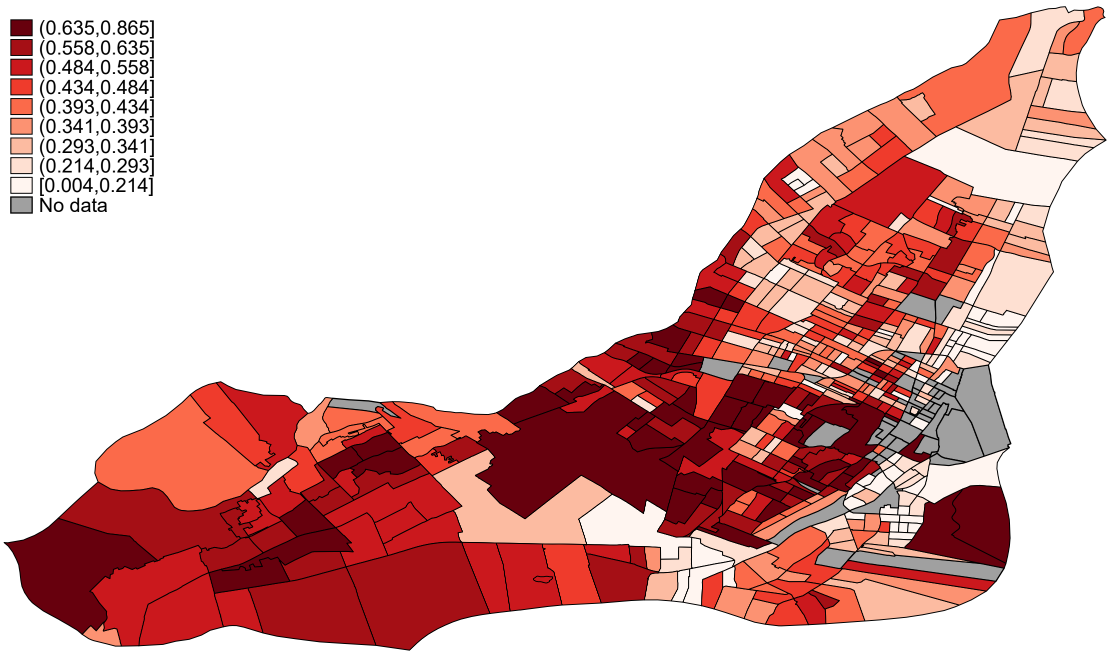
Figure A6: Fraction in Private Schools, by Grade and Language of Instruction



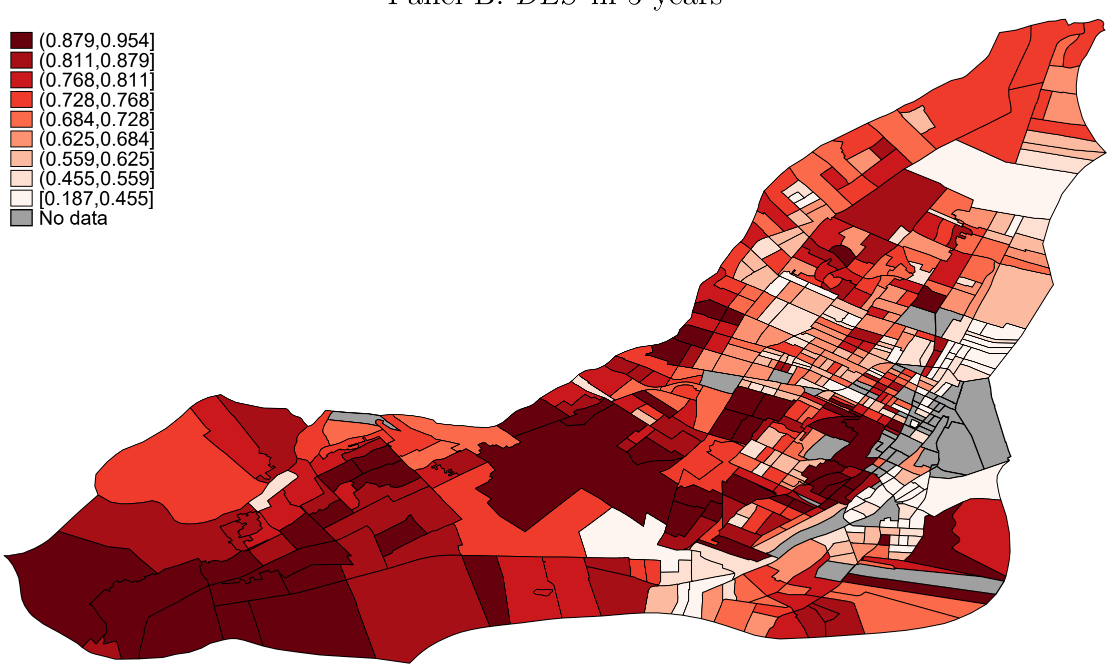
Notes: This figure shows the fraction of students enrolled in a private school by grade. Statistics calculated over main analytical sample of 92,764 students who completed their compulsory schooling in Montreal. Data shown separately for students in French and English schools.

Figure A7: Spatial Variation in Educational Outcomes - Census Tract Level

Panel A: University enrollment



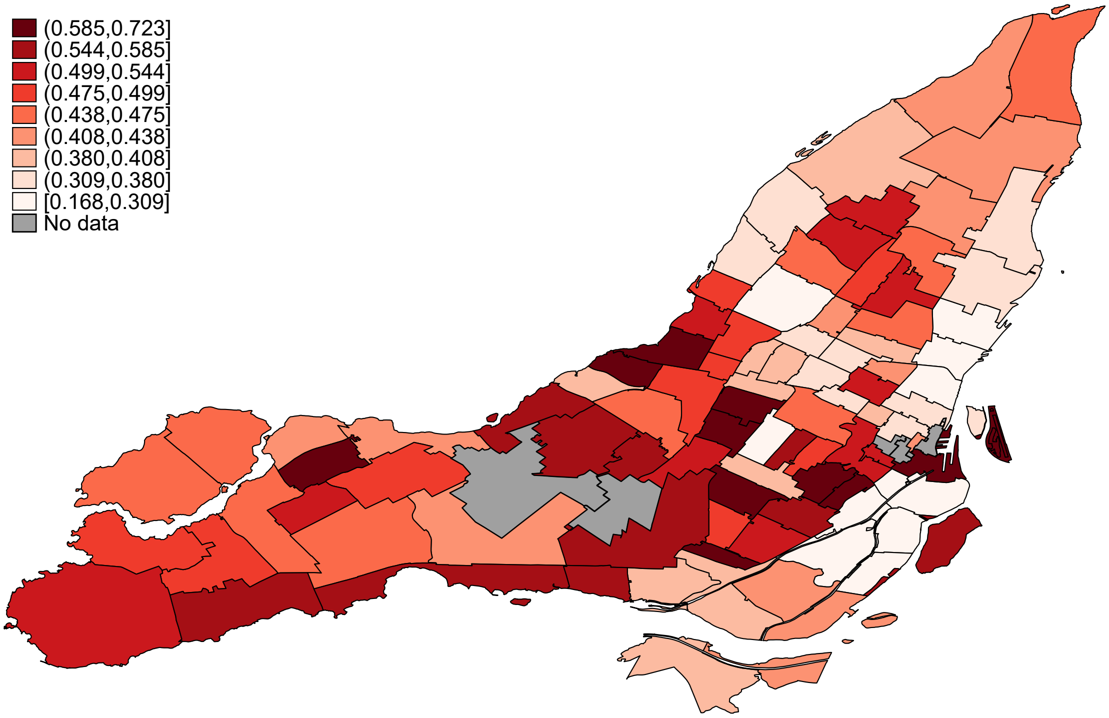
Panel B: *DES* in 5 years



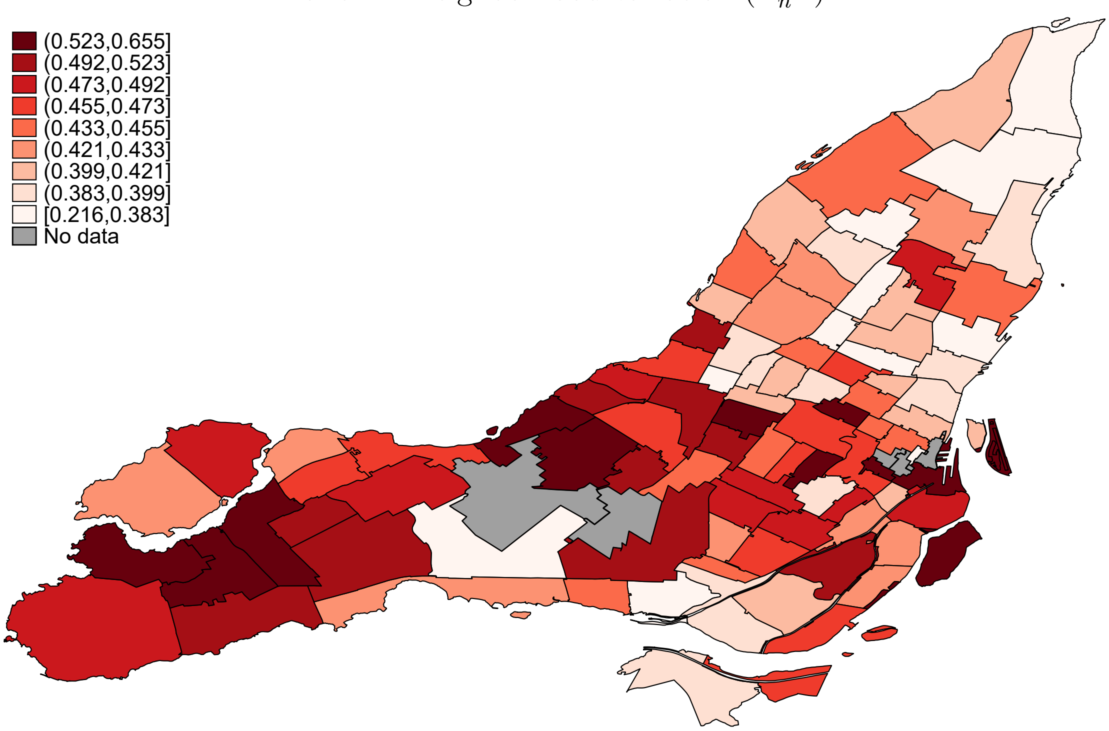
Notes: These maps replicates Figure 1, but average educational attainment is now measured at the census tract level. The sample is restricted to students who always resided in the same census tract (permanent residents). Census tracts with fewer than 10 permanent residents are in grey (labeled "No data").

Figure A8: Spatial Variation in $\bar{\Omega}_n^{PR}$ and Λ_n^{PR}, for University Enrollment

Panel A: School variation ($\bar{\Omega}_n^{PR}$)

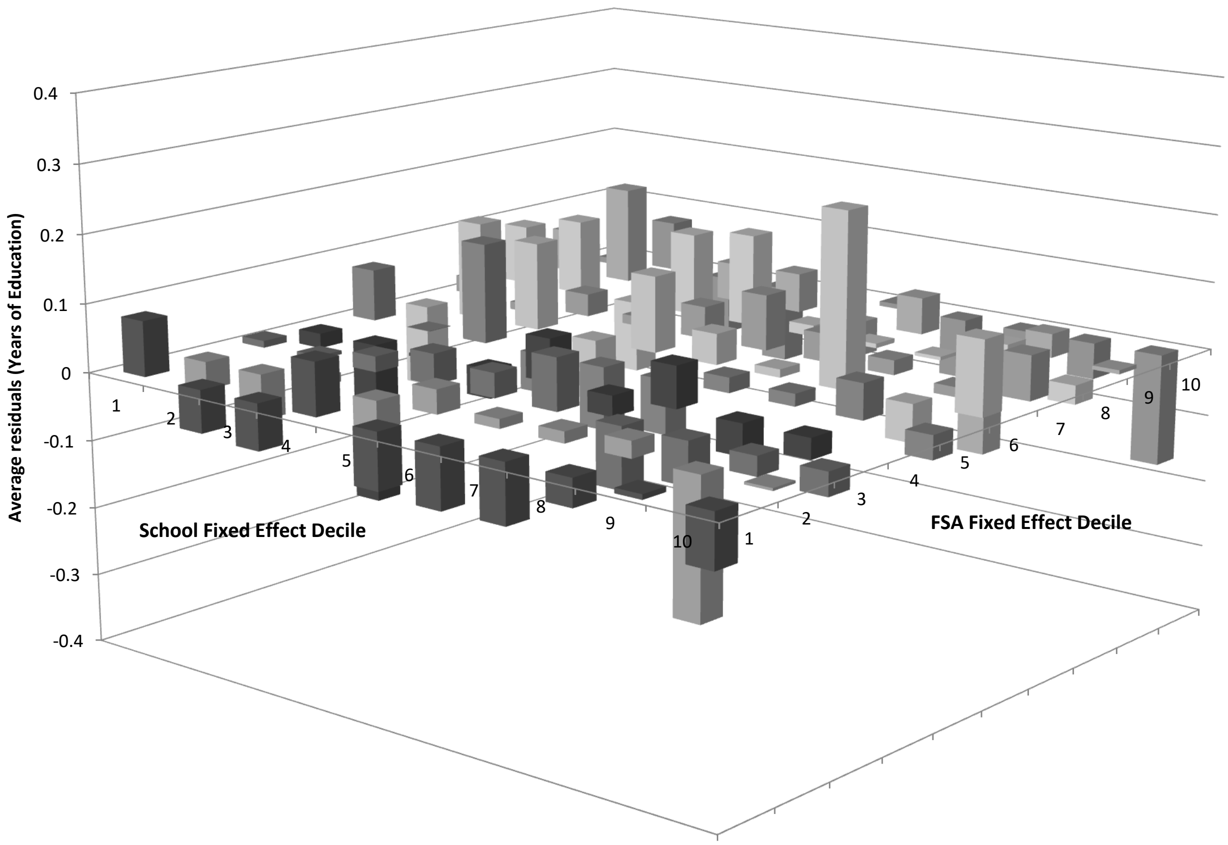


Panel B: Neighborhood variation (Λ_n^{PR})



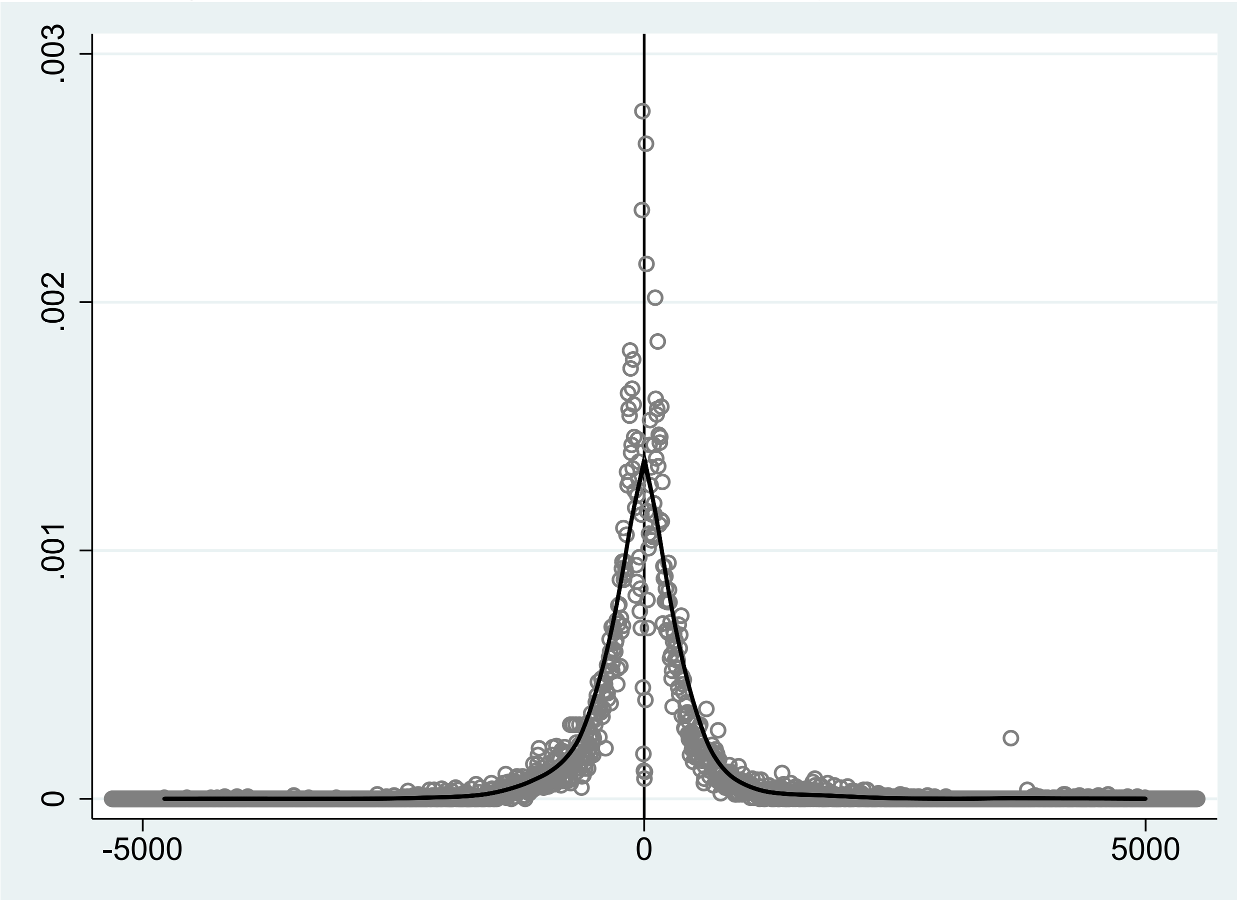
Notes: These maps plot FSA-level averages of school (panel A) and neighborhood (panel B) fixed effects estimated from equation (1). In constructing these estimates, the sample is restricted to permanent residents. The measure of educational attainment is university enrollment, which is equal to one for students who were ever enrolled in a Quebec university, and zero otherwise. To ease the interpretation, the values of $\bar{\Omega}_n^{PR}$ and Λ_n^{PR} are re-centered around the unconditional university enrollment rate for the sample of permanent residents. Data for FSAs with fewer than 10 permanent residents are in grey (labeled "No data").

Figure A9: Mean Years of Education (Residuals), by School and FSA Deciles



Notes: This figures plots average residuals from a regression of years of education on school, neighborhood and cohort fixed effects (equation (1)). The figure is constructed by slicing the student-level distributions of school and FSA fixed effects into deciles, and then calculating the average residuals in each school-by-neighborhood decile cell. The estimation sample is restricted to permanent residents.

Figure A10: Density Plot around French Primary School Boundaries

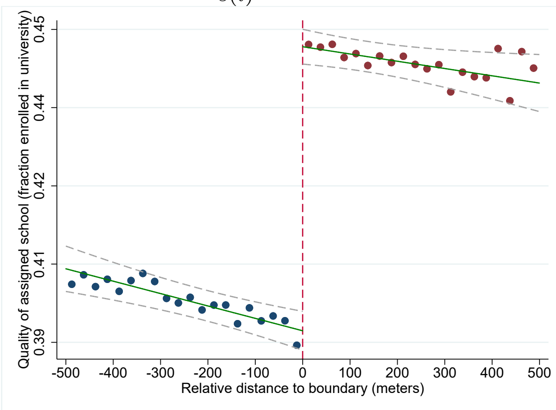


Notes: This figure shows a density plot of number of students by distance to the nearest French primary school boundary, in meters. For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}_s^P$ are at positive distances. Figure produced with the Stata package DCdensity.ado, which implements the test derived in McCrary (2008).

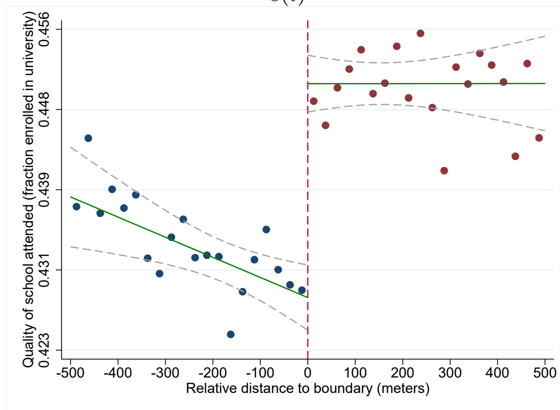
Figure A11: Discontinuities in School Quality at French Primary School Boundaries

All permanent residents

Panel A: Quality $\hat{\Omega}^P_{s(i)}$ of assigned French school

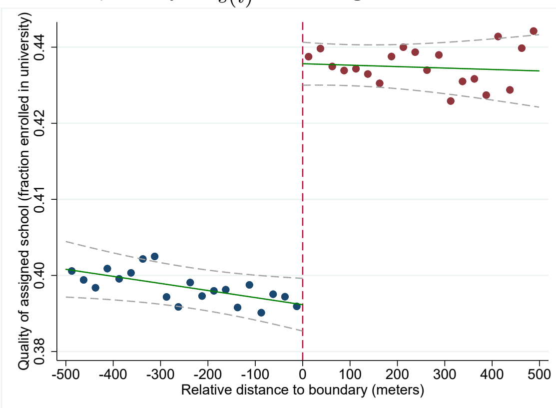


Panel B: Quality $\hat{\Omega}^P_{s(i)}$ of school attended

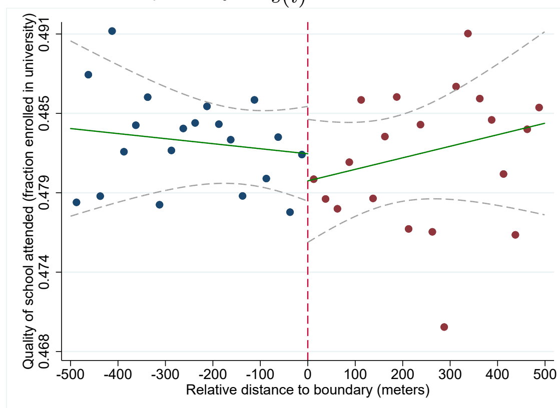


Students in English schools only (Placebo)

Panel C: Quality $\hat{\Omega}^P_{s(i)}$ of assigned French school

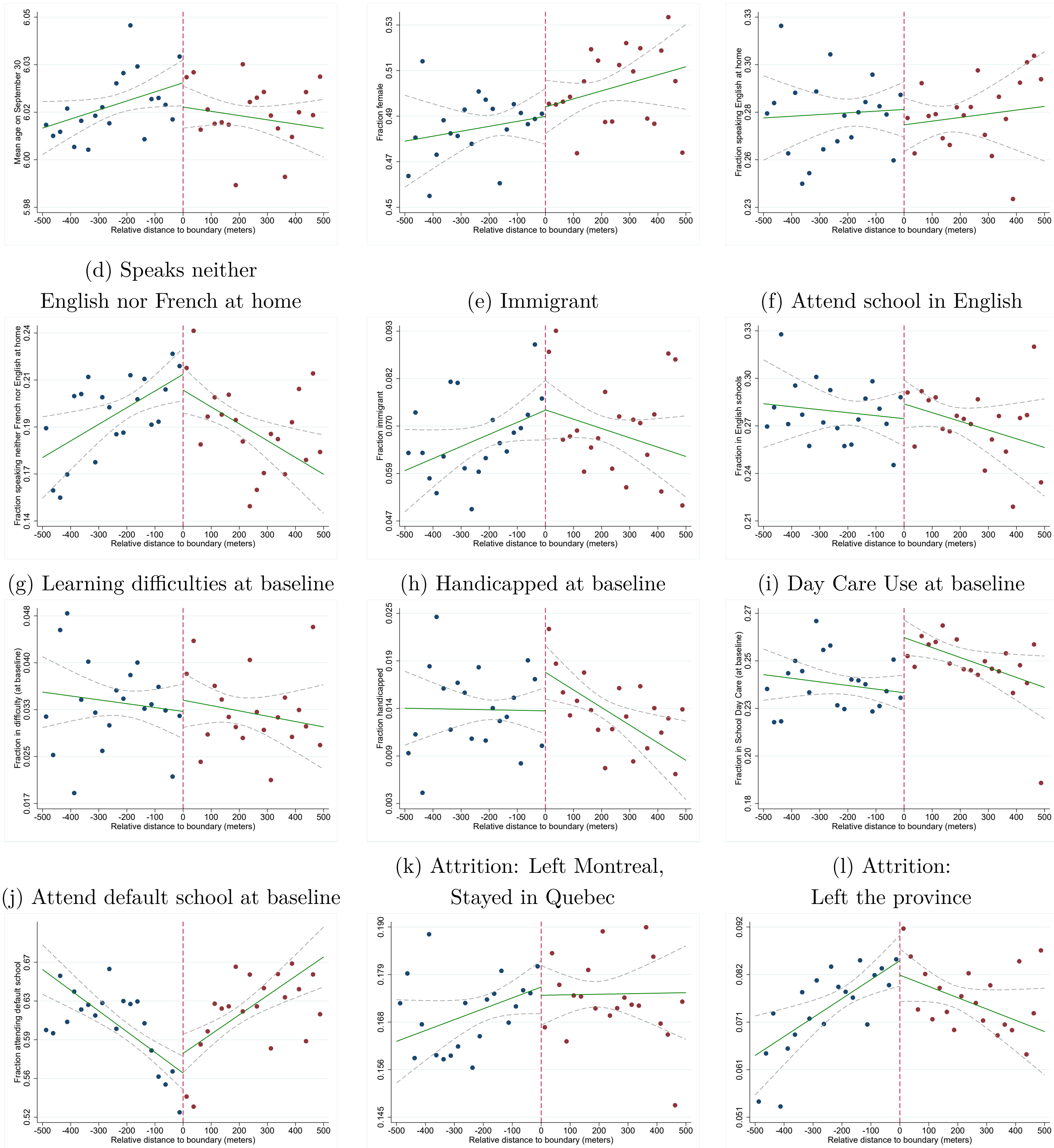


Panel D: Quality $\hat{\Omega}^P_{s(i)}$ of school attended



Notes: This figure presents the relationship between quality of the default French primary school (panels A and C) and distance to the nearest French primary school boundary, and between quality of the primary school attested and distance to the boundary (panels B and D). For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}^P_s$ (in therm of university enrollment) are at positive distances. Variables on the vertical axis are residualized on cohort, FSA, and boundary fixed effects. In Panels A and B, the sample includes all permanent residents, and in Panels C and D it is restricted to students enrolled in English schools. For visual clarity, students living further than 500 meters away from their nearest boundary are excluded. Each dot indicates average values of the (residualized) dependent variable within 25-meter bins. Solid green lines are linear fits of the relationship between the dependent variable and distance to the nearest boundary. Grey dashed lines are 95% confidence intervals based on standard errors clustered at the boundary level.

Figure A12: Balance of Covariates at Boundaries - School Quality in Terms of University Enrollment



(a) Age (b) Gender (c) Speaks English at home

(d) Speaks neither English nor French at home (e) Immigrant (f) Attend school in English

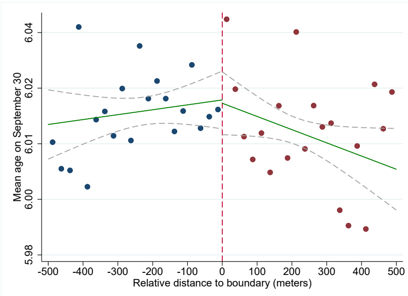
(g) Learning difficulties at baseline (h) Handicapped at baseline (i) Day Care Use at baseline

(j) Attend default school at baseline (k) Attrition: Left Montreal, Stayed in Quebec (l) Attrition: Left the province

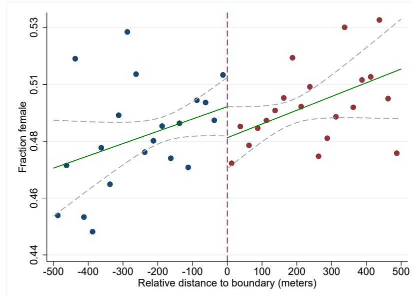
Notes: This figure presents the relationship between student characteristics and distance to the nearest French primary school boundary. In panels (a) to (j), the sample is restricted to permanent residents. In panels (k) and (l), there is no sample restriction, hence all students in the database are included. For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}_s^P$ (measured in rates of university enrollment) are at positive distances. For visual clarity, students living further than 500 meters away from their nearest boundary are excluded. Each dot indicates average values of the (residualized) dependent variable within 25-meter bins. Solid green lines are linear fits of the relationship between the dependent variable and distance to the nearest boundary. Grey dashed lines are 95% confidence intervals based on standard errors clustered at the boundary level.

Figure A13: Balance of Covariates at Boundaries - School Quality in Terms of *DES* in 5 Years

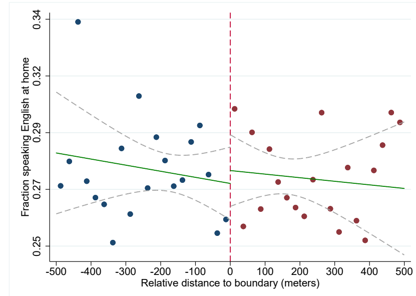
(a) Age



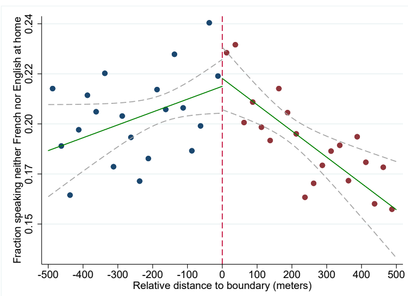
(b) Gender



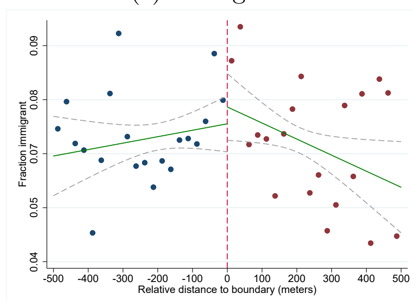
(c) Speaks English at home



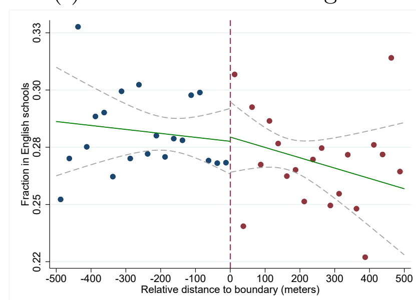
(d) Speaks neither English nor French at home



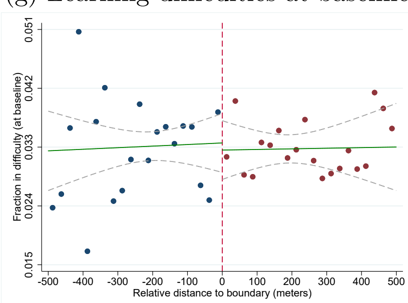
(e) Immigrant



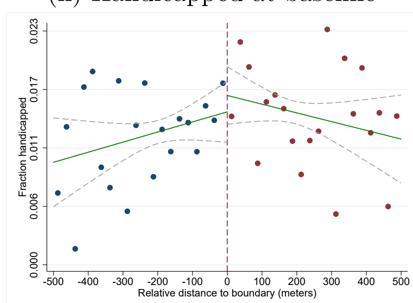
(f) Attend school in English



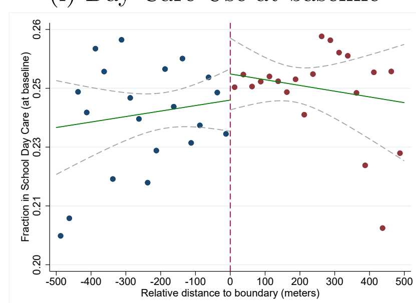
(g) Learning difficulties at baseline



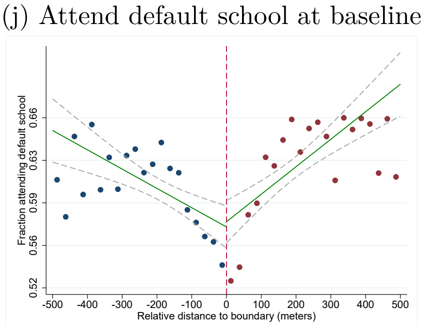
(h) Handicapped at baseline



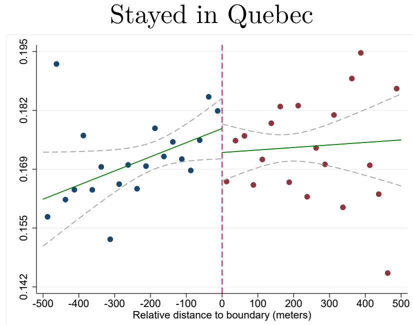
(i) Day Care Use at baseline



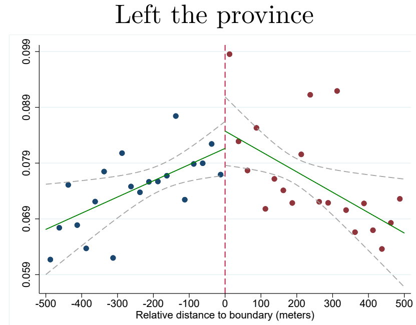
(j) Attend default school at baseline



(k) Attrition: Left Montreal, Stayed in Quebec

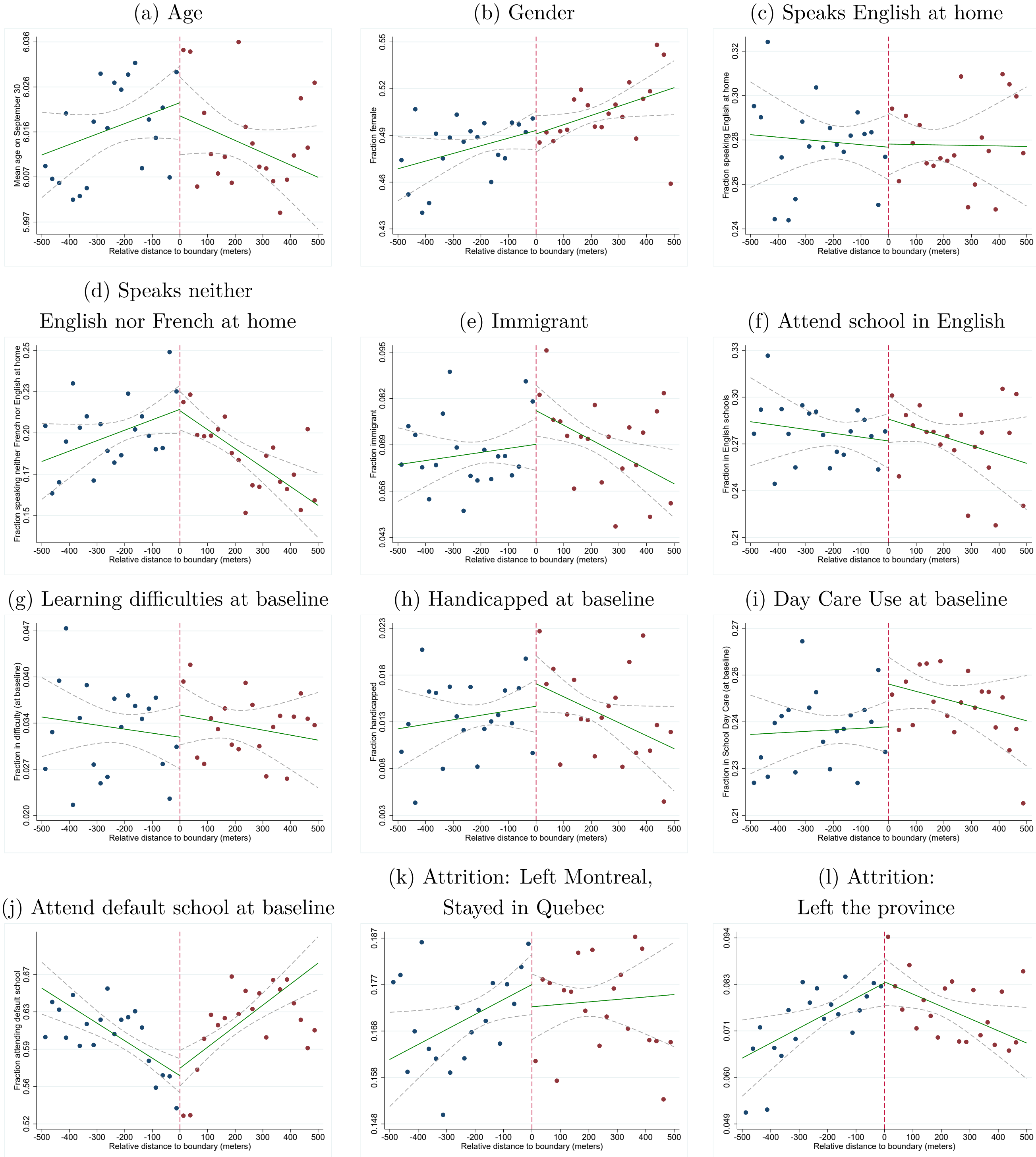


(l) Attrition: Left the province



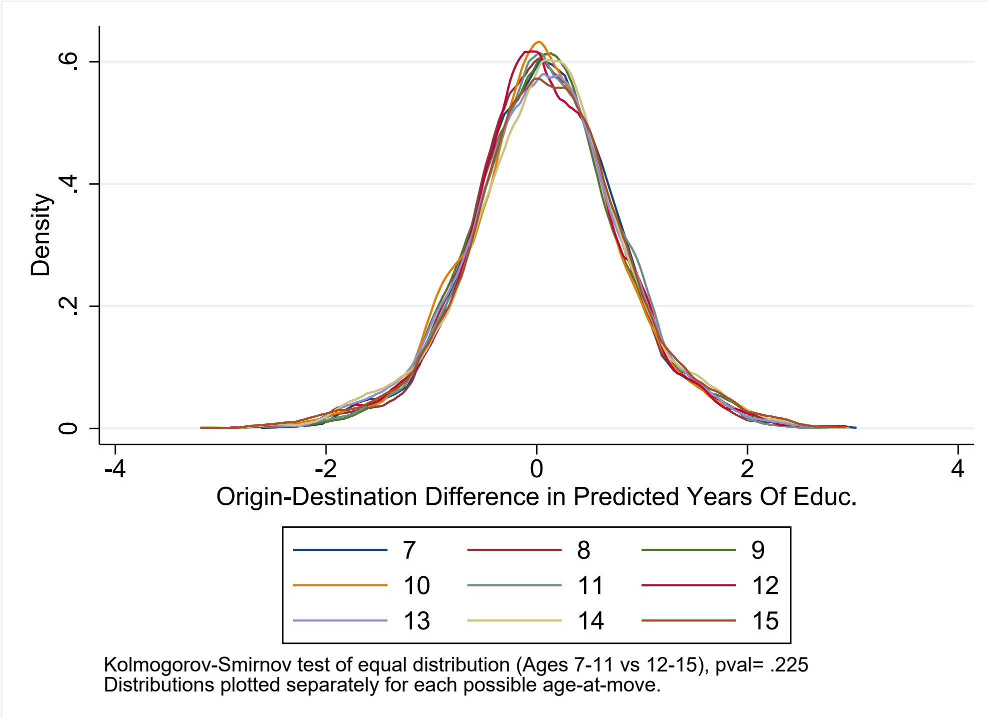
Notes: This figure presents the relationship between student characteristics and distance to the nearest French primary school boundary. In panels (a) to (j), the sample is restricted to permanent residents. In panels (k) and (l), there is no sample restriction, hence all students in the database are included. For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}_s^P$ (measured in rates of timely secondary school graduation) are at positive distances. For visual clarity, students living further than 500 meters away from their nearest boundary are excluded. Each dot indicates average values of the (residualized) dependent variable within 25-meter bins. Solid green lines are linear fits of the relationship between the dependent variable and distance to the nearest boundary. Grey dashed lines are 95% confidence intervals based on standard errors clustered at the boundary level.

Figure A14: Balance of Covariates at Boundaries - School Quality in Terms of Years of Education



Notes: This figure presents the relationship between student characteristics and distance to the nearest French primary school boundary. In panels (a) to (j), the sample is restricted to permanent residents. In panels (k) and (l), there is no sample restriction, hence all students in the database are included. For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}_s^P$ (measured in years of education) are at positive distances. For visual clarity, students living further than 500 meters away from their nearest boundary are excluded. Each dot indicates average values of the (residualized) dependent variable within 25-meter bins. Solid green lines are linear fits of the relationship between the dependent variable and distance to the nearest boundary. Grey dashed lines are 95% confidence intervals based on standard errors clustered at the boundary level.

Figure A15: Distribution of $\Delta\bar{y}_{od}$ by Age-at-move

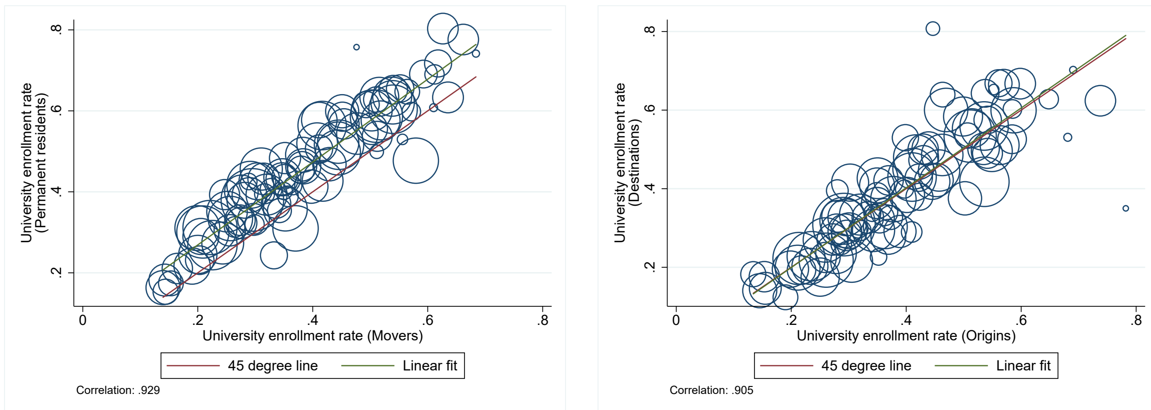


Notes: This figure shows kernel density plots of the distribution of $\Delta\bar{y}_{od}$ – the difference in years of education between permanent residents of destination d and origin o – for each possible value of age-at-move m. It also reports the p-value for a Kolmogorov-Sminov test of equality of distribution for moves at ages 7-11 and moves at ages 12-15. The sample is restricted to movers within Montreal.

Figure A16: FSA-level Mean Outcomes of Permanent Residents and of Movers

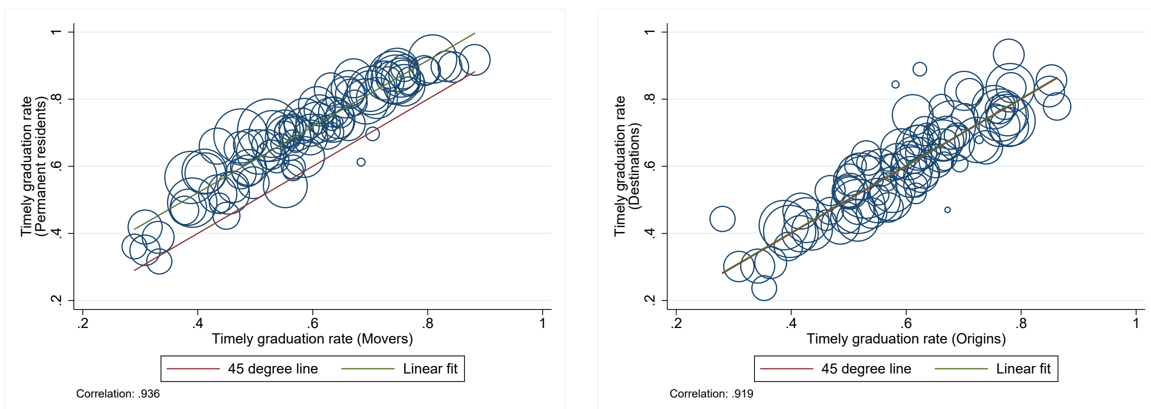
Panel A: University enrollment

Movers vs Permanent Residents | Movers to vs Movers from



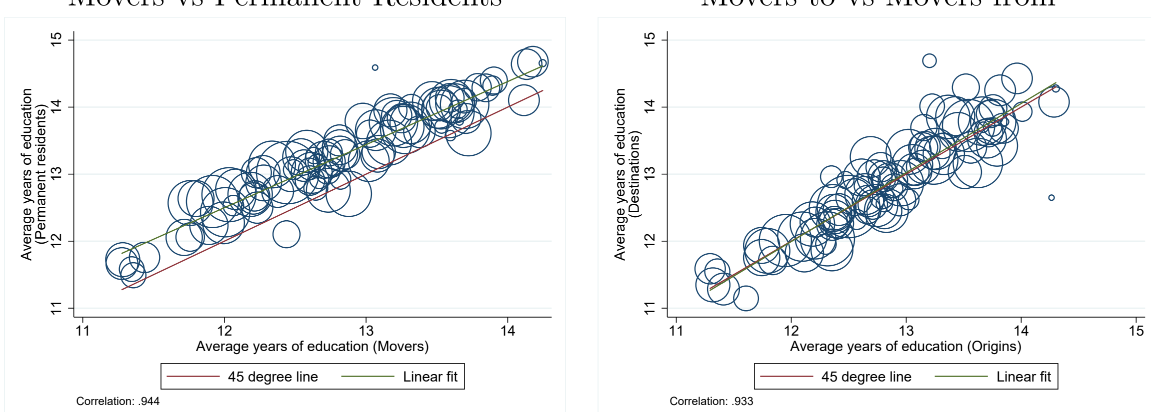
Panel B: DES in 5 Years

Movers vs Permanent Residents | Movers to vs Movers from



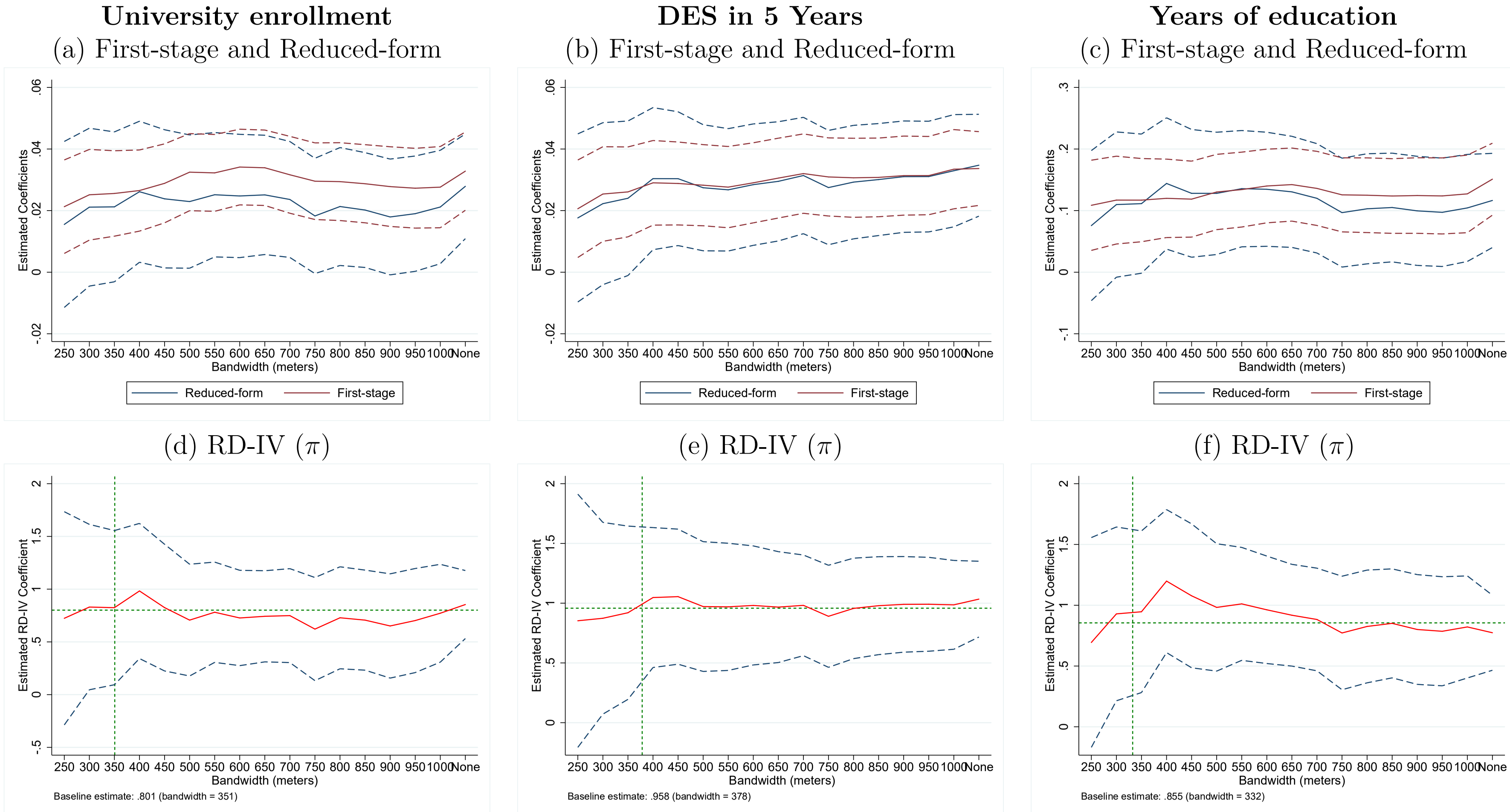
Panel C: Years of education

Movers vs Permanent Residents | Movers to vs Movers from



Notes: Figures on the left plot the FSA-level average educational attainment of permanent residents (y-axis) against the exposure-weighted average for movers (x-axis). For movers, if child i has lived 5 years in FSA a and 5 years in FSA b, that child receives a weight of 50% in each of these two FSAs. Figures on the right plot exposure-weighted average outcomes for movers using only years spent in their destinations (y-axis) against exposure-weighted average outcomes for these same movers using only years spent in their origin. That is, the same mover is used to calculate the average outcome of 2 different FSAs. In all figures, each circle represents one FSA, and circle size is proportional to the number of permanent resident in each FSA. Mean outcomes are residualized on cohort fixed effects. Data for FSAs with fewer than 10 permanent residents are not shown. The red line is the 45 degree line, and the green line is the linear fit of the y-axis variable as a function of the x-axis variable.

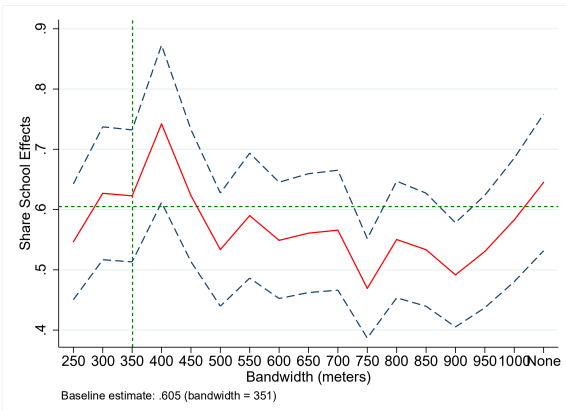
Figure A17: Bandwidth Sensitivity of RD Estimates



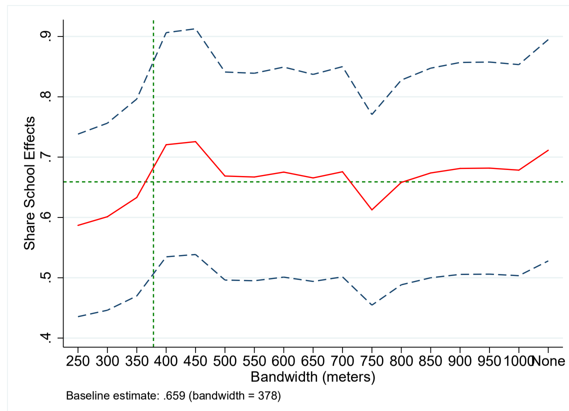
Notes: Panels (a) to (c) show first-stage and reduced-form RD estimates of the reduced-form effects of residing on the higher school quality side of a French primary school boundary for different bandwidth values. The dashed blue lines represent 95% confidence intervals with standard errors clustered at the boundary-level. Panels (d) to (f) show the associated RD-IV coefficients of the causal effect of school quality on educational attainment. The horizontal dashed green line shows the baseline RD-IV estimate reported in Table 2 and based on Calonico, Cattaneo and Titiunik (2014) optimal bandwidths. The vertical dashed green line shows the value of the baseline optimal bandwidth.

Figure A18: Bandwidth Sensitivity of School Share S^{school}

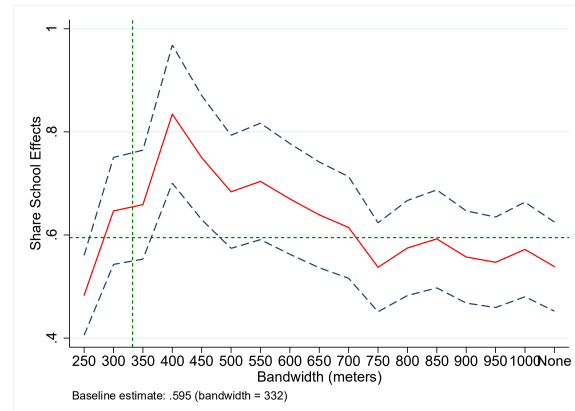
Panel A: University enrollment



Panel B: DES in 5 Years



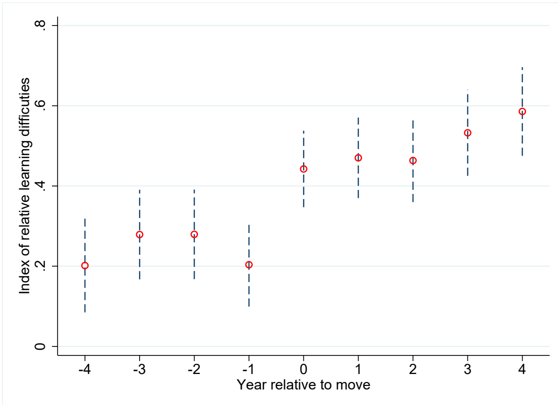
Panel C: Years of education



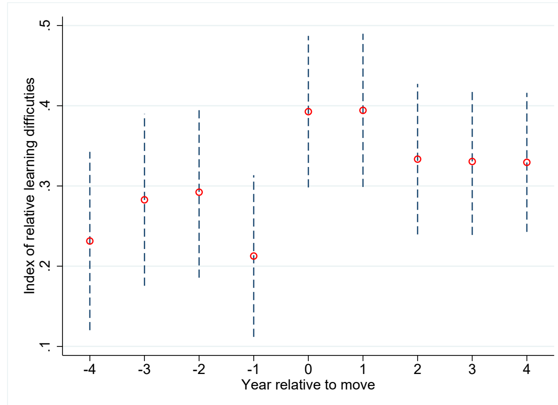
Notes: This figure shows estimates of the share of total exposure effects that is due to school quality S^{school} for different RD bandwidth values. The red line show point estimates, and the dashed blue lines represent 95% confidence intervals with standard errors calculated by the delta method. The horizontal dashed green line shows the baseline estimate of the school share S^{school} obtained using Calonico et al. (2014) optimal bandwidth. The vertical dashed green line shows the value of the baseline optimal bandwidth.

Figure A19: Index of Relative Learning Difficulties, by Years Relative to Move

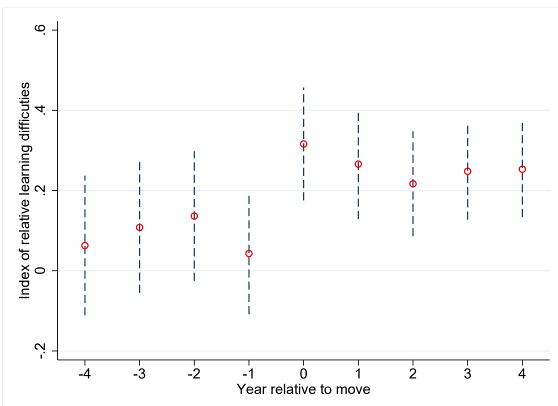
Panel A: No student fixed effects



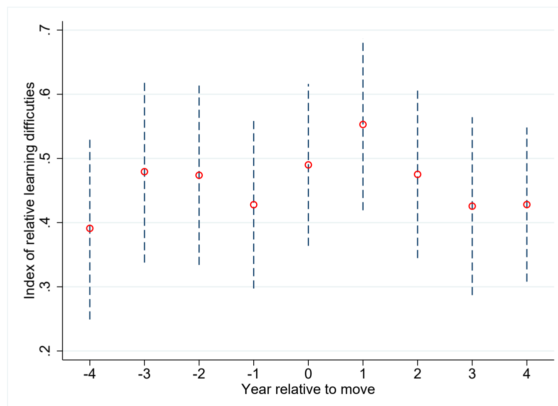
Panel B: With student fixed effects



Panel C: With student fixed effects
School switchers



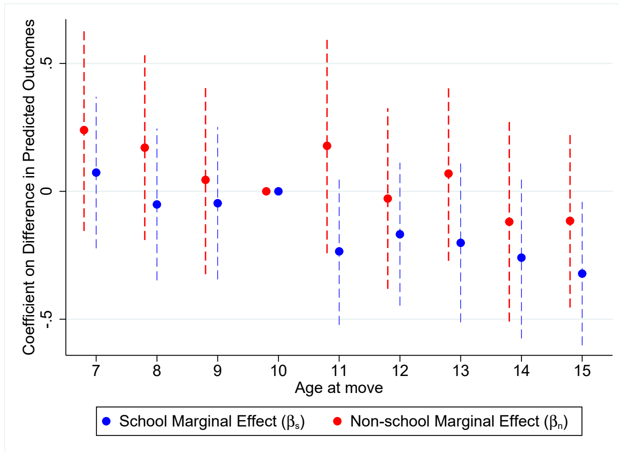
Panel D: With student fixed effects
Non-switchers



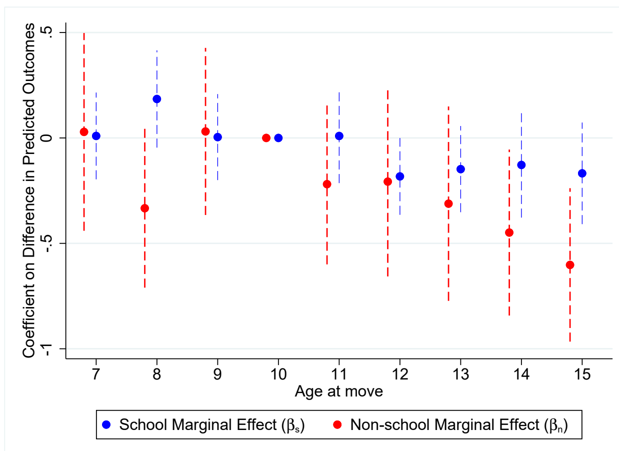
Notes: This figure shows event-studies of student learning difficulties around the time of a residential move. The y-axis shows regression coefficients of $\sigma_{od(i,t)} = \frac{Diff_{iod,t} - \overline{Diff}_{o,t}}{\overline{Diff}_{d,t} - \overline{Diff}_{o,t}}$ – an index of relative learning difficulties that summarizes the way movers compare to permanent residents in their origin and destination – on relative-time dummies, net of cohort, age and time (i.e. years since started grade 1) fixed effects. Observations outside the event window are included in the regression, so all coefficients are relative to omitted relative-time periods. In panel A and B, all movers are included in the estimating sample. Panel C includes only students who switched school the year they moved. Panel D includes movers who did not switch school the year they moved. Observations are weighted by $\left(\overline{Diff}_{dt} - \overline{Diff}_{ot}\right)^2$, and weights are normalized within time periods (i.e. weights sum to one within each period). Dashed lines represent 95% confidence intervals with standard errors clustered at the student level.

Figure A20: Semi-parametric Partial Exposure Effects

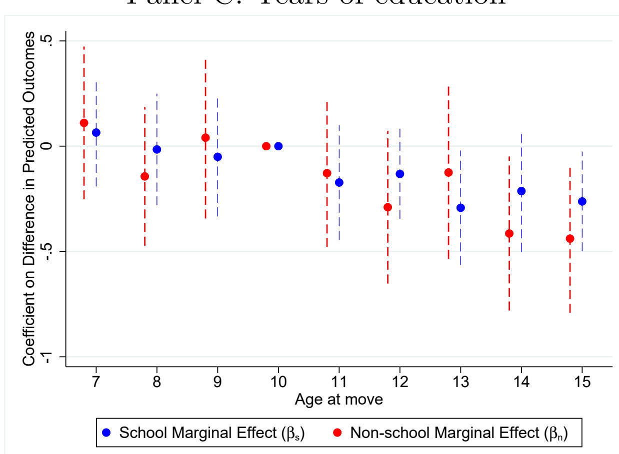
Panel A: University enrollment



Panel B: *DES* in 5 Years



Panel C: Years of education

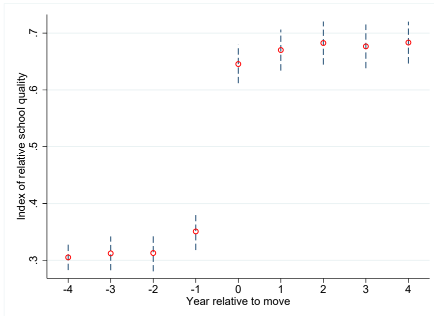


Notes: This figure presents regression estimates of partial exposure effects on educational attainment. Blue dots correspond to age-specific partial regression coefficients of y_i on $\pi\Delta\Omega_{od}$, and red dots are similarly defined for coefficients on $\Delta\bar{y}_{od}^{-s}$ in the following regression: $y_{icmod} = \sum_{m=7}^{15} \beta_{s,m} \left(\pi\Delta\Omega_{od} \times 1\{m_i = m\}\right) + \sum_{m=7}^{15} \beta_{n,m} \left(\Delta\bar{y}_{od}^{-s} \times 1\{m_i = m\}\right) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod}$, where i indexes a student, o the origin FSA, d the destination FSA, and c cohorts. In panel A, the outcome y is university enrollment, in panel B it is an indicator for finishing secondary school within 5 years, and in panel C it is the number of years of education. All coefficients are relative to moving at age 10. The sample includes all movers who remained within Montreal. Students who moved from or to FSAs with less than 10 permanent residents are excluded. Standard errors are clustered at the destination FSA level. Dashed vertical lines show 95% confidence intervals for the point estimates.

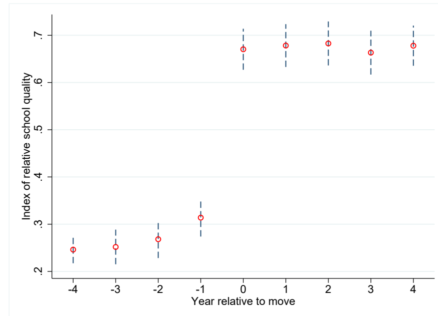
Figure A21: Index of Relative School Quality, by Years Relative to Move

Panel A: University enrollment

All movers



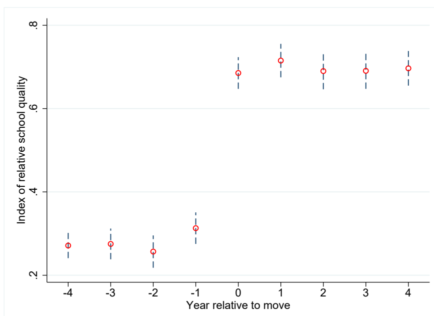
One-time movers



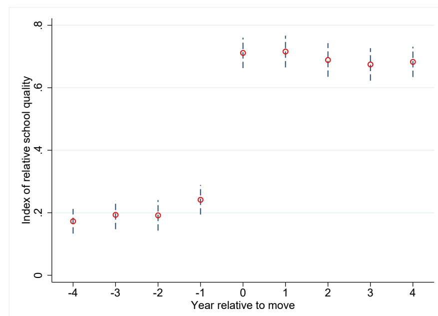
Pre-post difference: 0.47 (s.e. 0.014) Pre-post difference: 0.54 (s.e. 0.018)

Panel B: DES in 5 Years

All movers



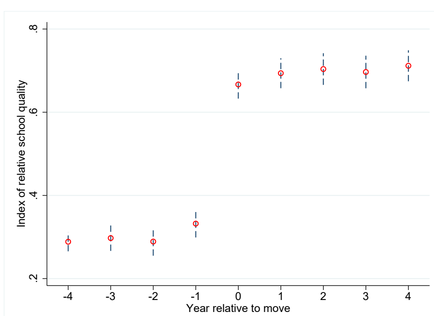
One-time movers



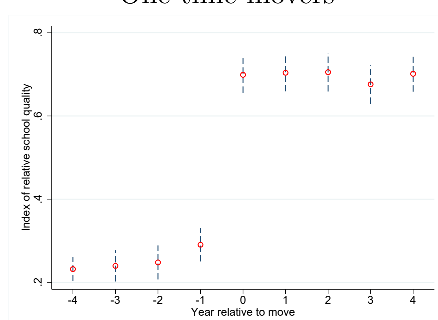
Pre-post difference: 0.56 (s.e. 0.019) Pre-post difference: 0.65 (s.e. 0.024)

Panel C: Years of education

All movers



One-time movers



Pre-post difference: 0.52 (s.e. 0.015) Pre-post difference: 0.59 (s.e. 0.019)

Notes: This figure shows event-studies of student learning difficulties around the time of a residential move. The y-axis shows regression coefficients of $\sigma^{\psi}_{od(i,t)} = \frac{\Omega_{s(i,t)} - \overline{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}$ – an index of relative school quality that summarizes the way school attended by movers compare to those attended by permanent residents in their origin and destination – on relative-time dummies, net of cohort, age and time (i.e. years since started grade 1) fixed effects. For each period t, $\overline{\Omega}_{s(n,t)}$ is measured by the relevant average primary school fixed effects if student i was in primary school in that year. Secondary school fixed effects are used for remaining years. Observations outside the event window are included in the regression, so all coefficients are relative to omitted relative-time periods. In panel A and B, all movers are included in the estimating sample. Panel C includes only students who switched school the year they moved. Panel D includes movers who did not switch school the year they moved. Observations are weighted by $\left(\overline{Diff}_{dt} - \overline{Diff}_{ot}\right)^2$, and weights are normalized within time periods (i.e. weights sum to one within each period). Dashed lines represent 95% confidence intervals with standard errors clustered at the student level.

Table A1: Descriptive Statistics

	All students	Permanent residents	Movers: Within Montreal	Movers: Left Montreal	Difference between (2) and (3)
	mean	mean	mean	mean	coef.
	(s.d.)	(s.d.)	(s.d.)	(s.d.)	(s.e.)
Variables	(1)	(2)	(3)	(4)	(5)
Female	0.49	0.49	0.49	0.49	0.001
	[0.500]	[0.500]	[0.500]	[0.500]	(0.004)
Age on September 30	6.02	6.01	6.04	6.02	-0.033
	[0.378]	[0.329]	[0.457]	[0.333]	(0.003)
Mother tongue: French	0.49	0.47	0.44	0.64	0.026
	[0.500]	[0.499]	[0.497]	[0.479]	(0.004)
Mother tongue: English	0.21	0.26	0.20	0.09	0.063
	[0.408]	[0.439]	[0.399]	[0.293]	(0.003)
Mother tongue: Other	0.30	0.27	0.36	0.26	-0.090
	[0.458]	[0.444]	[0.480]	[0.440]	(0.003)
Language at home: French	0.53	0.50	0.50	0.69	0.006
	[0.499]	[0.500]	[0.500]	[0.462]	(0.004)
Language at home: English	0.26	0.32	0.24	0.12	0.077
	[0.437]	[0.466]	[0.428]	[0.327]	(0.003)
Language at home: Other	0.21	0.18	0.26	0.19	-0.083
	[0.407]	[0.384]	[0.441]	[0.388]	(0.003)
Immigrant	0.10	0.07	0.14	0.11	-0.080
	[0.300]	[0.247]	[0.352]	[0.310]	(0.002)
Language at school: French	0.75	0.69	0.77	0.88	-0.072
	[0.433]	[0.461]	[0.424]	[0.325]	(0.003)
Uses School Day Care (baseline)	0.25	0.24	0.24	0.29	0.005
	[0.432]	[0.428]	[0.425]	[0.452]	(0.003)
In difficulty (baseline)	0.04	0.03	0.05	0.05	-0.016
	[0.194]	[0.170]	[0.210]	[0.219]	(0.001)
Handicapped (baseline)	0.01	0.01	0.02	0.01	-0.002
	[0.118]	[0.116]	[0.126]	[0.111]	(0.001)
Ever in difficulty by age 15	0.31	0.25	0.36	0.37	-0.116
	[0.461]	[0.431]	[0.481]	[0.483]	(0.003)
Students	92,764	44,912	31,526	16,326	76,438

Notes: This table reports descriptive statistics for the main sample, which excludes students who left Quebec's system before turning 16. Permanent residents are defined as students who always resided in the same FSA until the age of 15. Movers within Montreal are those who moved across FSAs at least once and were still living on the Island of Montreal at age 15. Movers who left Montreal were residing in the province of Quebec but outside the Island of Montreal at age 15. For columns (1) through (4), standard deviations are in square brackets. Column (5) reports the difference in means between permanent residents and movers within Montreal, and the associated standard errors are in brackets.

Table A2: Summary Statistics: Educational Outcomes Across Cohorts

		Cohort				
	All	**1995**	**1996**	**1998**	**2000**	**2001**
Primary and secondary school outcomes						
Did not start secondary school on time	0.113	0.156	0.153	0.124	0.073	0.068
Secondary school diploma	0.760	0.755	0.752	0.759	0.767	0.765
Secondary school diploma in 5 years	0.610	0.600	0.587	0.609	0.630	0.621
No secondary school qualification	0.200	0.208	0.209	0.195	0.189	0.198
Post-secondary outcomes						
Ever enrolled in college	0.695	0.678	0.682	0.699	0.710	0.705
Enrolled in college by age 17	0.530	0.497	0.503	0.532	0.560	0.555
Ever enrolled in university	0.373	0.460	0.451	0.424	0.332	0.220
Enrolled in university by age 19	0.170	0.166	0.166	0.169	0.175	0.175
Bachelor degree or more	0.128	0.275	0.249	0.140	0.003	0.004
Educational attainment						
Number of years of education	12.810	13.247	13.200	13.066	12.517	12.119
Observations	92,764	16,969	18,067	18,777	19,125	19,826

Notes: This table shows cohort-specific average educational attainment for the main sample, which excludes students who left Quebec's system before turning 16. These statistics exclude almost 1,000 individuals who enroll in a Quebec post-secondary institution at some point, but who had left the primary and secondary school system before turning 16 and therefore are excluded from the main sample.

Table A3: Selected Census Tract Characteristics: Montreal, Boston and the US

	Greater Montreal (2001)				Greater Boston (2000)			
	Overall	Census-tract level variation			Overall	Census-tract level variation		
	mean	sd	p5	p95	mean	sd	p5	p95
% families w children single-headed	0.280	0.129	0.125	0.529	0.241	0.167	0.086	0.643
% Black	0.041	0.051	0.000	0.142	0.065	0.157	0.002	0.452
% Foreign born	0.194	0.161	0.020	0.504	0.149	0.115	0.029	0.382
Employment rate	0.608	0.090	0.463	0.748	0.642	0.086	0.479	0.745
% Below poverty line	0.222	0.144	0.058	0.510	0.086	0.098	0.016	0.304
% Some college	0.479	0.160	0.278	0.797	0.623	0.180	0.301	0.895
# of tracts	850				701			
Total population	3,426,350				3,406,829			
	Chicago (2000)				U.S. (2000)			
	Overall	Census-tract level variation			Overall	Census-tract level variation		
	mean	sd	p5	p95	mean	sd	p5	p95
% families w children single-headed	0.259	0.220	0.080	0.772	0.282	0.165	0.101	0.654
% Black	0.186	0.366	0.002	0.980	0.121	0.235	0.001	0.782
% Foreign born	0.172	0.146	0.001	0.452	0.111	0.132	0.003	0.405
Employment rate	0.617	0.137	0.333	0.767	0.597	0.115	0.386	0.746
% Below poverty line	0.105	0.145	0.013	0.420	0.124	0.117	0.019	0.373
% Some college	0.568	0.212	0.200	0.890	0.518	0.190	0.220	0.839
# of tracts	1,864				65,174			
Total population	8,272,768				281,421,906			

Notes: This table shows mean demographics for the CMA of Montreal, for the PMSA of Boston, the PMSA of Chicago, and for the US as a whole. Data for Montreal is from the 2001 Canadian Census. Data for Boston, Chicago and the US is from the 2000 US Census and was extracted via The National Historical Geographic Information System (Steven Manson and Ruggles, 2019). For each location, the first column report the sample means of each characteristics. The next three columns are all based on census tract-level data. It reports the between-tract standard deviation, as well as the 5th and 95th percentile of the distribution across tracts. For Montreal, the % below the poverty line is the fraction of the population below Statistics Canada's Low-Income-Cut-Off.

Table A4: Variation Across FSAs and Schools - Empirical Bayes Estimates

	Outcome					
	University enrollment		DES in 5 years		Years of education	
	(1)	(2)	(3)	(4)	(5)	(6)
Student-level standard deviation of shrunk fixed effects:						
Neighborhoods (FSAs)	0.129	0.035	0.126	0.016	0.633	0.150
Schools	0.237	0.218	0.259	0.251	1.156	1.073
Dependent variable summary statistics:						
Mean	0.443		0.706		13.228	
Standard deviation	[0.497]		[0.456]		[2.113]	
Fixed effects estimated						
Separately	x		x		x	
Simultaneously		x		x		x
Number of students	44,912					
Number of primary schools	440					
Number of secondary schools	218					
Number of neighborhoods	95					

Notes: This table reports shrunk estimates of variance components of educational attainment of students who always resided in the same FSA (permanent residents). It replicates Table 1, but reports the standard deviation of shrunk estimates rather than unadjusted fixed effect estimates. To shrink estimates, I first calculate standard errors for each school and neighborhood fixed effect using bootstrap resampling (300 samples with replacement, clustering within primary school-secondary school-FSA cells). I then shrink estimates toward their means using the empirical Bayes procedure described in Chandra et al. (2016). Note that the reported empirical Bayes measures of school effects are shrunk estimates of the sum of primary and secondary school effects, not the sum of shrunk estimates of primary school and shrunk estimates of secondary school effects.

Table A5: Balance of Covariates at Boundaries

Outcome used to assign *HighSide:*	University Enrollment	DES in 5 years	Years of education
	(1)	(2)	(3)
Covariates			
Age	-0.0076	0.0033	0.0058
	(0.0094)	(0.0090)	(0.0096)
Gender	0.0000	-0.0143	-0.0020
	(0.0118)	(0.0109)	(0.0126)
Speaks English at Home	-0.0135	0.0120	0.0096
	(0.0135)	(0.0132)	(0.0136)
Speaks neither French nor English at Home	0.0065	0.0044	0.0027
	(0.0123)	(0.0119)	(0.0127)
Immigrant	0.0030	0.0066	0.0141
	(0.0075)	(0.0072)	(0.0078)
Attend school in English	0.0051	0.0108	0.0170
	(0.0144)	(0.0141)	(0.0142)
Learning difficulties at baseline	0.0046	0.0007	0.0053
	(0.0044)	(0.0044)	(0.0045)
Handicapped at baseline	0.0060	0.0007	0.0039
	(0.0031)	(0.0030)	(0.0032)
Day Care Use at baseline	0.0334	0.0068	0.0154
	(0.0080)	(0.0082)	(0.0089)
Attend default school at baseline	0.0317	0.0011	0.0036
	(0.0185)	(0.0181)	(0.0192)
Left Montreal	-0.0077	-0.0098	-0.0131
	(0.0074)	(0.0070)	(0.0075)
Left the province	0.0016	0.0042	0.0059
	(0.0052)	(0.0052)	(0.0053)
Predicted educational attainment	-0.0053	-0.0009	-0.0148
	(0.0024)	(0.0036)	(0.0149)
Cohort fixed effects	x	x	x
Individual characteristics	x	x	x
Neighborhood (FSA) fixed effects	x	x	x
Boundary fixed effects	x	x	x

Notes: This table shows estimates of discontinuities in the distribution of covariates at boundaries. These are coefficients θ from the estimating equation $X_{icnb} = \theta HighSide_b + f(distance_{ib}) + \alpha_b + \alpha_n + \alpha_c + \epsilon_{icnb}$, where $HighSide_b$ is an indicator for residing on the higher school quality side of French primary school boundary b. In all specifications, the control function for distance to boundary is linear and allows for different slopes on either side of the threshold. Regression discontinuity bandwidths are the same as in the baseline specification (Table 2). The underlying sample includes all permanent residents, expect for the attrition variables (Left Montreal and Left the province), where all students in the database are included. Predicted educational attainment is given by the fitted values of a regression of the outcome of interest on gender, place of birth indicators, language at home indicators, use of day care, "in difficulty" status at baseline, handicapped status and cohort fixed effects. All standard errors are clustered at the French primary school boundary level.

Table A6: Exposure Effects: Alternative Outcomes

Sample:	All movers		One-time movers	
	(1)	(2)	(3)	(4)
Measure of educational attainment				
No Secondary school qualification	-0.0676	-0.0648	-0.0496	-0.0496
	(0.0137)	(0.0143)	(0.0159)	(0.0165)
College enrollment (ever)	-0.0373	-0.0356	-0.0267	-0.0258
	(0.0109)	(0.0118)	(0.0134)	(0.0138)
College enrollment by 17	-0.0412	-0.0382	-0.0408	-0.0389
	(0.0081)	(0.0081)	(0.0112)	(0.0114)
College degree	-0.0407	-0.0389	-0.0336	-0.0321
	(0.0087)	(0.0090)	(0.0110)	(0.0111)
University enrollment by 19	-0.0395	-0.0381	-0.0454	-0.0442
	(0.0110)	(0.0112)	(0.0160)	(0.0163)
Bachelor degree or more	-0.0374	-0.0363	-0.0261	-0.0258
	(0.0129)	(0.0130)	(0.0181)	(0.0182)
Expected earnings on basis of	-0.0455	-0.0433	-0.0411	-0.0397
level of education	(0.0087)	(0.0093)	(0.0097)	(0.0097)
Expected earnings on basis of	-0.0407	-0.0392	-0.0334	-0.0324
level and field of education	(0.0085)	(0.0090)	(0.0106)	(0.0105)
Cohort fixed effects	x	x	x	x
Individual characteristics	x	x	x	x
Age at move fixed effects	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x
Only moved once			x	x
Times in difficulty before moving		x		x
N	24316	24316	15533	15533

Notes: This table reports regression estimates of total exposure effects on educational attainment. It replicates Table 3, but for alternative measured of educational attainment. The movers sample contains a total of 25,993 observations, of which 1,677 are singletons and therefore dropped in the estimation. Details on the measurement of outcomes are provided in the Data Appendix. See notes to Table 3 for additional details.

Table A7: Total Exposure Effects: Exposure-weighted Neighborhood Quality for Multiple-times Movers

Sample:	All movers			
	(1)	(2)	(3)	(4)
Measure of educational attainment				
University enrollment	-0.0534	-0.0511	-0.0497	-0.0484
	(0.0096)	(0.0096)	(0.0096)	(0.0097)
Secondary school diploma in 5 years	-0.0515	-0.0483	-0.0466	-0.0451
	(0.0089)	(0.0089)	(0.0090)	(0.0092)
Years of education	-0.0607	-0.0582	-0.0563	-0.0550
	(0.0095)	(0.0099)	(0.0098)	(0.0102)
N	24316	24316	24191	24191
Cohort fixed effects	x	x	x	x
Individual characteristics	x	x	x	x
Age at move fixed effects	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x
Number of moves fixed effects	x	x	x	x
Other locations controls			x	x
Times in difficulty before moving to *d*		x		x

Notes: This table reports regression estimates of total exposure effects on educational attainment. It replicates Table 3, but accounts for multiple moves. The sample includes all movers in Montreal. The change in neighborhood quality is measured by $\bar{y}_d - E(\bar{y}_n|premove)$, where $E(\bar{y}_n|premove)$ is the exposure-weighted average neighborhood quality for all locations in which student i resided before moving to the final destination d. Note that $\bar{y}_d - E(\bar{y}_n|premove) = \Delta\bar{y}_{od}$ for one-time movers. All specifications include dummies for number of moves before the age of 15. In columns (3) and (4), fixed effects for the second and third location (prior to moving to area d), as well as for the age at which these moves occurred, are included (the omitted categories are no second/third location). Standard errors are clustered at the final destination neighborhood level. See notes to Table 3 for additional details.

Table A8: Total Exposure Effects: 6-digit Postal Code Fixed Effects

Sample:	All movers			One-time movers		
	(1)	(2)	(3)	(4)	(5)	(6)
Measure of educational attainment						
University enrollment	-0.0424	-0.0412	-0.0403	-0.0416	-0.0408	-0.0538
	(0.0090)	(0.0092)	(0.0117)	(0.0116)	(0.0115)	(0.0214)
Secondary school diploma in 5 years	-0.0421	-0.0402	-0.0443	-0.0506	-0.0502	-0.0301
	(0.0088)	(0.0088)	(0.0151)	(0.0117)	(0.0117)	(0.0227)
Years of education	-0.0488	-0.0471	-0.0462	-0.0444	-0.0435	-0.0409
	(0.0088)	(0.0094)	(0.0116)	(0.0103)	(0.0102)	(0.0189)
Cohort fixed effects	x	x	x	x	x	x
Individual characteristics	x	x	x	x	x	x
Age at move fixed effects	x	x	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x	x	x
Only moved once				x	x	x
Times in difficulty before moving		x	x		x	x
Destination 6-digit postal code fixed effects			x			x
N	24316	24316	16525	15533	15533	8856

Notes: This table reports regression estimates of total exposure effects on educational attainment. Columns (1), (2), (4) and (5) replicate results presented in Table 3. Columns (3) and (6) additionally control for 6-digit postal code fixed effects at age 15. Sample sizes are considerably lower in these specifications because of the larger number of singletons. Standard errors are clustered at the destination neighborhood level. See notes to Table 3 for additional details.

Table A9: Total Exposure Effects: Heterogeneous Exposure Effects

Heterogeneity by:	Gender		Language at school		Moves to	
	Boys	Girls	French	English	Better FSA	Worse FSA
	(1)	(2)	(3)	(4)	(5)	(6)
Measure of educational attainment						
University enrollment	-0.0321	-0.0571	-0.0385	-0.039	-0.0398	-0.0536
	(0.0123)	(0.0128)	(0.0107)	(0.0207)	(0.0192)	(0.0207)
Secondary school diploma in 5 years	-0.044	-0.0476	-0.0473	-0.0449	-0.0385	-0.0115
	(0.0110)	(0.0140)	(0.0114)	(0.0205)	(0.0137)	(0.0235)
Years of education	-0.0425	-0.0587	-0.0485	-0.0525	-0.0257	-0.052
	(0.0119)	(0.0127)	(0.0124)	(0.0173)	(0.0151)	(0.0192)
Cohort fixed effects	x	x	x	x	x	x
Individual characteristics	x	x	x	x	x	x
Age at move fixed effects	x	x	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x	x	x
N	11600	11283	17479	5832	10981	13335

Notes: This table documents heterogeneity in estimates of total exposure effects on educational attainment. It replicates Table 3, but separately for subsamples of students. Column (1) includes only boys and column (2) restricts the sample to girls. In columns (3) and (4), regressions are run separately by language of instruction at age 15. In column (5), the sample is restricted to movers for which $\Delta\bar{y}_{od} > 0$ and column (6) is restricted to cases where $\Delta\bar{y}_{od} < 0$. Standard errors are clustered at the destination neighborhood level. See notes to Table 3 for additional details.

Table A10: School Effects: Quadratic Control Function

	First-stage(s)			Reduced-form	RD-IV (π)	Parameters	
	Dependent variable:						
	Quality of *assigned* school at baseline ($\Omega^{P}_{s(i)}$)	Quality of school *attended* at baseline ($\Omega^{P}_{s(i)}$)	Childhood average school quality ($\Omega^{-i}_{s(i,n)}$)	Outcome	Outcome	Bandwidth	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Measure of educational attainment							
			All permanent residents				
University enrollment	0.0634	0.0248	0.0293	0.0206	0.7086		
	(0.0032)	(0.0028)	(0.0065)	(0.0093)	(0.2235)		
Secondary school diploma in 5 years	0.0714	0.0304	0.0325	0.0351	1.0812	Global	43291
	(0.0036)	(0.0026)	(0.0063)	(0.0091)	(0.1891)		
Years of education	0.2961	0.1192	0.1372	0.1098	0.8023		
	(0.0146)	(0.0126)	(0.0309)	(0.0428)	(0.1914)		
			Placebo: Students in English schools				
University enrollment	0.0624	-0.0051	-0.0089	-0.0169	-		
	(0.0043)	(0.0041)	(0.0106)	(0.0178)			
Secondary school diploma in 5 years	0.0712	0.0056	0.0012	-0.0052	-	Global	13444
	(0.0054)	(0.0028)	(0.0077)	(0.0135)			
Years of education	0.2810	0.0042	-0.0075	-0.0438	-		
	(0.0189)	(0.0167)	(0.0478)	(0.0770)			
Cohort fixed effects	x	x	x	x	x		
Individual characteristics	x	x	x	x	x		
Neighborhood (FSA) fixed effects	x	x	x	x	x		
Boundary fixed effects	x	x	x	x	x		

Notes: This table replicates Table 2, changing the functional form of the control function for distance to boundary in the regression discontinuity model to a global quadratic polynomial. It allows for different functions on either side of the threshold. In the first three rows, the sample includes all permanent residents. In the last three rows, only permanent residents enrolled in English schools are included. All standard errors are clustered at the French primary school boundary level. See notes to Table 2 for additional details.

Table A11: School Effects: Triangular Kernel Control Function

	First-stage(s)			Reduced-form	RD-IV (π)	Parameters	
	Dependent variable:						
	Quality of *assigned* school at baseline ($\Omega^{P}_{s(i)}$)	Quality of school *attended* at baseline ($\Omega^{P}_{s(i)}$)	Childhood average school quality ($\Omega^{-i}_{s(i,n)}$)	Outcome	Outcome	Bandwidth (meters)	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Measure of educational attainment							
	All permanent residents						
University enrollment	0.0639	0.0226	0.0235	0.0197	0.8313	401	32133
	(0.0033)	(0.0028)	(0.0070)	(0.0121)	(0.4015)		
Secondary school diploma in 5 years	0.0701	0.0289	0.0261	0.0259	0.9925	470	34183
	(0.0039)	(0.0027)	(0.0070)	(0.0119)	(0.3495)		
Years of education	0.2906	0.0943	0.1048	0.0899	0.8538	312	28205
	(0.0154)	(0.0133)	(0.0369)	(0.0609)	(0.4352)		
	Placebo: Students in English schools						
University enrollment	0.0617	-0.0011	0.0022	-0.0144	-	376	8304
	(0.0047)	(0.0040)	(0.0102)	(0.0213)			
Secondary school diploma in 5 years	0.0704	0.0062	0.0045	-0.0013	-	417	8760
	(0.0059)	(0.0035)	(0.0086)	(0.0197)			
Years of education	0.2738	0.0111	0.0339	-0.0814	-	349	7904
	(0.0208)	(0.0161)	(0.0448)	(0.0956)			
Cohort fixed effects	x	x	x	x	x		
Individual characteristics	x	x	x	x	x		
Neighborhood (FSA) fixed effects	x	x	x	x	x		
Boundary fixed effects	x	x	x	x	x		

Notes: This table replicates Table 2, using a triangular kernel for the control function for distance to boundary in the regression discontinuity model. It allows for different functions on either side of the threshold. In the first three rows, the sample includes all permanent residents. In the last three rows, only permanent residents enrolled in English schools are included. All standard errors are clustered at the French primary school boundary level. See notes to Table 2 for additional details.

Table A12: School Effects: Re-weighted Sample

	First-stage	Reduced-form	RD-IV
Dependent variable:	Childhood average school quality ($\Omega^{-i}_{s(l,n)}$)	Outcome	Outcome
	(1)	(2)	(3)
Measure of educational attainment			
University enrollment	0.0339	0.026	0.7706
	(0.0061)	(0.0083)	(0.1670)
Secondary school diploma in 5 years	0.0345	0.0351	1.0189
	(0.0059)	(0.0089)	(0.1837)
Years of education	0.1557	0.1095	0.705
	(0.0286)	(0.0376)	(0.1580)
N	43287	43292	43287
Cohort fixed effects	x	x	x
Individual characteristics	x	x	x
Neighborhood (FSA) fixed effects	x	x	x
Boundary fixed effects	x	x	x

Notes: This table present regression discontinuity estimates of the causal effect of school quality on educational attainment, where the full sample of permanent residents is re-weighted so that its distribution of covariates (gender, place of birth indicators, language at home indicators, use of day care, "in difficulty" status at baseline, handicapped status) matches that of the movers' sample used in Table 3. The matching weights are obtained using nearest-neighbor matching (5 nearest neighbors) with the Stata command `kmatch` (Jann, 2017). In all specifications, the econometric specification is identical to that described in Table 2 except that to insure balance of covariates between the two sample, there is no bandwidth restriction in these analyses. All standard errors are clustered at the French primary school boundary level. See notes to Table 2 for additional details.

Table A13: Total Exposure Effects: Balancing Tests

Outcome of permanent residents:	University Enrollment		DES in 5 years		Years of Education	
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates						
Gender	0.0187	0.0162	0.0175	0.0174	0.0038	0.0034
	(0.0095)	(0.0138)	(0.0093)	(0.0136)	(0.0019)	(0.0029)
Speaks English at Home	-0.0121	-0.0034	-0.0149	-0.0035	-0.0024	-0.0001
	(0.0084)	(0.0113)	(0.0076)	(0.0107)	(0.0016)	(0.0021)
Speaks neither French nor English at Home	-0.0030	-0.0044	0.0033	0.0036	-0.0006	-0.0010
	(0.0081)	(0.0107)	(0.0077)	(0.0093)	(0.0015)	(0.0020)
Immigrant	-0.0175	-0.0246	-0.0054	-0.0104	-0.0032	-0.0045
	(0.0065)	(0.0098)	(0.0069)	(0.0095)	(0.0013)	(0.0019)
Handicapped	-0.0036	-0.0049	-0.0039	-0.0037	-0.0006	-0.0008
	(0.0027)	(0.0036)	(0.0027)	(0.0039)	(0.0006)	(0.0008)
Use Day Care at baseline	0.0026	-0.0016	-0.0004	-0.0021	0.0003	-0.0005
	(0.0066)	(0.0078)	(0.0069)	(0.0081)	(0.0014)	(0.0016)
In difficulty at baseline	0.0078	0.0085	0.0066	0.0052	0.0017	0.0016
	(0.0035)	(0.0043)	(0.0036)	(0.0045)	(0.0007)	(0.0009)
Times in difficulty pre-move	0.0365	0.0391	0.0313	0.0125	0.0098	0.0078
	(0.0394)	(0.0407)	(0.0394)	(0.0399)	(0.0081)	(0.0083)
Cohort fixed effects	x	x	x	x	x	x
Age at move fixed effects	x	x	x	x	x	x
Origin-by-Destination fixed effects	x	x	x	x	x	x
One-time movers only		x		x		x
N	24316	15533	24316	15533	24316	15533

Notes: This table presents estimates of the relationship between individual characteristics and the interaction between age-at-move and the difference in outcomes between permanent residents of the destination and origin neighborhoods. It reports estimates of b in estimating equation $X_{icmod} = b\,(m_i \times \Delta\bar{y}_{od}) + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod}$, where i indexes a student, o the origin FSA, d the destination FSA, and c cohorts. In Columns (1) and (2), $\Delta\bar{y}_{od}$ is measured using years of education. In columns (3) and (4), fractions of students finishing secondary school in 5 years are used, and in columns (5) and (6), university enrollment rates are. The sample includes all movers in Montreal. Standard errors are clustered at the destination neighborhood level.

Table A14: Total Exposure Effects: Robustness to Time-varying Observables

	(1)	(2)	(3)	(4)	(5)	(6)
Measure of educational attainment						
University enrollment	-0.0373	-0.0399	-0.0436	-0.0392	-0.0350	-0.0332
	(0.0107)	(0.0107)	(0.0095)	(0.0097)	(0.0110)	(0.0117)
Secondary school diploma in 5 years	-0.0392	-0.0380	-0.0437	-0.0423	-0.0370	-0.0318
	(0.0101)	(0.0105)	(0.0087)	(0.0093)	(0.0100)	(0.0111)
Years of education	-0.0435	-0.0457	-0.0484	-0.0437	-0.0422	-0.0376
	(0.0096)	(0.0095)	(0.0090)	(0.0090)	(0.0103)	(0.0105)
Time-varying controls						
Income	x					x
Percent low-income		x				x
Dwelling value			x			x
Percent lone family				x		x
Percent with college					x	x
Cohort fixed effects	x	x	x	x	x	x
Individual characteristics	x	x	x	x	x	x
Age at move fixed effects	x	x	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x	x	x
N	22735	22735	22735	22735	22735	22735

Notes: This table examines the robustness of regression estimates of total exposure effects on educational attainment to the inclusion of time-varying control variables. It replicates Table 3, but adds differences in census tract characteristics ΔZ_{iod} for student i around the time of the move, as well as their interaction with age-at-move, as control variables. The table presents estimates of β from estimating equation $y_{icmod} = \beta\left(m_i \times \Delta \bar{y}_{od}\right) + \eta_0 \Delta Z_{iod} + \eta_1\left(m_i \times \Delta Z_{iod}\right) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod}$, Census tract characteristics are obtained from the 2001 Canadian Census. The model includes both the main effect of these controls as well as their interaction with age-at-move. Each column includes a different set of observable time-varying variables ΔZ_{iod}. Standard errors are clustered at the destination neighborhood level. See notes to Table 3 for additional details on the movers design.

Table A15: Decomposition of Exposure Effects: Measurement Error

Outcome:	University enrollment		DES in 5 years		Years of education	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Empirical Bayes						
			Total exposure effects			
β (PR outcomes = y^{EB})	-0.0450	-0.0450	-0.0466	-0.0466	-0.0523	-0.0523
	(0.0097)	(0.0097)	(0.0098)	(0.0098)	(0.0094)	(0.0094)
β (PR outcomes = $\Omega^{EB} + \Lambda^{EB}$)	-0.0493	-0.0493	-0.0435	-0.0435	-0.0545	-0.0545
	(0.0105)	(0.0105)	(0.0103)	(0.0103)	(0.0102)	(0.0102)
			School and non-school components			
β^{school}	-0.0415	-0.0266	-0.0385	-0.0351	-0.0441	-0.0406
	(0.0091)	(0.0058)	(0.0102)	(0.0093)	(0.0094)	(0.0086)
$\beta^{non\text{-}school}$	-0.0078	-0.0227	-0.0050	-0.0084	-0.0104	-0.0139
	(0.0031)	(0.0052)	(0.0018)	(0.0020)	(0.0026)	(0.0028)
Share school effects (S^{school})	84%	54%	89%	81%	81%	75%
	(0.0513)	(0.0329)	(0.0470)	(0.0428)	(0.0455)	(0.0419)
Panel B: Split-Sample IV						
			Total exposure effects			
β	-0.0462	-0.0462	-0.0455	-0.0455	-0.0518	-0.0518
	(0.0096)	(0.0096)	(0.0087)	(0.0087)	(0.0089)	(0.0089)
			School and non-school components			
β^{school}	-0.0385	-0.0308	-0.0341	-0.0327	-0.0391	-0.0334
	(0.0081)	(0.0065)	(0.0084)	(0.0080)	(0.0078)	(0.0066)
$\beta^{non\text{-}school}$	-0.0077	-0.0154	-0.0113	-0.0127	-0.0127	-0.0184
	(0.0044)	(0.0049)	(0.0045)	(0.0044)	(0.0042)	(0.0044)
Share school effects (S^{school})	83%	67%	75%	72%	75%	64%
	(0.1762)	(0.1411)	(0.1843)	(0.1766)	(0.1498)	(0.1281)
π	1	RD estimate	1	RD estimate	1	RD estimate

Notes: This table reports regression estimates of total exposure effects β, the school component of exposure effects β^{school}, the non-school component of exposure effects$\beta^{non-school}$, and the school share of exposure effects S^{school}, as defined in the main text. It replicates Table 4, but using shrunk estimates of $\bar{y}_n^{EB}$, $\Omega_{s(i,n)}^{EB}$ and Λ_n^{EB}. To construct these shrunk estimates, I first calculate standard errors for each school and neighborhood fixed effect using bootstrap resampling (300 samples with replacement, clustering within primary school-secondary school-FSA cells). I then shrink estimates toward their means using the empirical Bayes procedure described in Chandra et al. (2016). Note that the empirical Bayes measures of school effects are shrunk estimates of the sum of primary and secondary school effects. In the first row, β is estimated via equation (4) using $m_i \times \Delta\bar{y}_{od}^{EB}$ as the main regressor, where $\Delta\bar{y}_{od}^{EB} = \bar{y}_d^{EB} - \bar{y}_o^{EB}$ and $\bar{y}_n^{EB}$ are empirical Bayes shrunk estimates of $\bar{y}_n$. In the second row, β is estimated using $m_i \times \Delta\left(\Omega_n^{EB} + \Lambda_n^{EB}\right)$ as the main regressor. The school and non-school components are estimated using $\pi\Omega_n^{EB}$ and $\Delta\bar{y}_{od}^{-s,EB} = (\Omega_n^{EB} + \Lambda_n^{EB}) - \pi\Omega_n^{EB}$. In column (1), π is set to one. In column (2), π is estimated using $\Omega_{s(i,n)}^{EB}$ in RD equations (2) and (3). Panel B presents estimates of a split-sample IV approach. In this approach, I first randomly split the sample of permanent residents in half, and calculate $\bar{y}_{n,g}$, $\bar{\Omega}_{n,g}$ and $\Lambda_{n,g}$ for each of the two samples $g \in \{1,2\}$. I then estimate the movers design using $\Delta\bar{y}_{od,1}$ as the main regressor, which I instrument with $\Delta\bar{y}_{od,2}$. Similarly, for decomposition equations, I instrument $\Delta\Omega_{od,1}$ with $\Delta\Omega_{od,2}$, and $\Delta\bar{y}_{od,1}^{-s}$ with $\Delta\bar{y}_{od,2}^{-s}$. In column (1), π is set to one, and in column (2) π is given by baseline RD estimates from Table 2. In all models, standard errors are clustered at the destination FSA level, and obtained by the delta method.

Table A16: Alternative Decomposition of Exposure Effects

Sample:	All movers			One-time movers		
	(1)	(2)	(3)	(4)	(5)	(6)
University enrollment						
			Total exposure effects			
β	-0.0424	-0.0424	-0.0424	-0.0416	-0.0416	-0.0416
	(0.0090)	(0.0090)	(0.0090)	(0.0116)	(0.0116)	(0.0116)
			School and non-school components			
β^{school}	-0.0322	-0.0321	-0.0264	-0.0322	-0.0321	-0.0269
	(0.0072)	(0.0072)	(0.0071)	(0.0095)	(0.0093)	(0.0098)
$\beta^{non\text{-}school}$	-0.0102	-0.0103	-0.0160	-0.0094	-0.0095	-0.0147
	(0.0041)	(0.0041)	(0.0065)	(0.0065)	(0.0065)	(0.0101)
Share school effects (S^{school})	76%	76%	62%	77%	77%	65%
	(0.0771)	(0.0768)	(0.1198)	(0.1304)	(0.1299)	(0.2004)
Secondary school diploma in 5 years						
			Total exposure effects			
β	-0.0421	-0.0421	-0.0421	-0.0506	-0.0506	-0.0506
	(0.0088)	(0.0088)	(0.0088)	(0.0117)	(0.0117)	(0.0117)
			School and non-school components			
β^{school}	-0.0305	-0.0286	-0.0264	-0.0397	-0.0370	-0.0350
	(0.0083)	(0.0081)	(0.0082)	(0.0109)	(0.0108)	(0.0110)
$\beta^{non\text{-}school}$	-0.0116	-0.0135	-0.0157	-0.0109	-0.0136	-0.0156
	(0.0039)	(0.0044)	(0.0051)	(0.0054)	(0.0057)	(0.0066)
Share school effects (S^{school})	72%	68%	63%	79%	73%	69%
	(0.0912)	(0.0986)	(0.1144)	(0.0989)	(0.1059)	(0.1210)
Years of education						
			Total exposure effects			
β	-0.0488	-0.0488	-0.0488	-0.0444	-0.0444	-0.0444
	(0.0088)	(0.0088)	(0.0088)	(0.0103)	(0.0103)	(0.0103)
			School and non-school components			
β^{school}	-0.0341	-0.0328	-0.0261	-0.0326	-0.0312	-0.0259
	(0.0081)	(0.0078)	(0.0081)	(0.0097)	(0.0094)	(0.0102)
$\beta^{non\text{-}school}$	-0.0147	-0.0159	-0.0227	-0.0118	-0.0132	-0.0185
	(0.0043)	(0.0045)	(0.0064)	(0.0060)	(0.0062)	(0.0087)
Share school effects (S^{school})	70%	67%	53%	73%	70%	58%
	(0.0834)	(0.0833)	(0.1187)	(0.1254)	(0.1268)	(0.1776)
Measure of school quality	$\pi\Omega_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$	$\pi\Omega_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$
π	1	1	RD estimate	1	1	RD estimate

Notes: This table replicates Table 4, using alternative definitions of the school and non-school components. Here, the parameters are defined as follow: $\beta^{school} = \beta_s \left(\frac{Var^r(\pi\Delta\Omega_{od})+Cov^r(\pi\Delta\Omega_{od},\Delta\bar{y}^{-s}_{od})}{Var^r(\Delta\bar{y}_{od})}\right)$, $\beta^{non-school} = \beta_n \left(\frac{Var^r(\Delta\bar{y}^{-s}_{od})+Cov^r(\pi\Delta\Omega_{od},\Delta\bar{y}^{-s}_{od})}{Var^r(\Delta\bar{y}_{od})}\right)$, and $S^{school} = \frac{\beta^{school}}{\beta}$. In columns (1) and (2), π is set to one. In column (3), π is given by the RD-IV estimates reported in Table 2. In columns (1) through (3), total exposure effects β correspond to estimates reported in Table 3, column (1). In columns (4) through (6), β correspond to estimates reported in Table 3 column (3). See notes to Table 4 for additional details.

B Data Appendix

Measurement of outcomes Different levels of education are governed by different departments of the Ministry of Education. Each department keeps separate student records in different formats, but these files can be matched using unique student IDs. Researchers interested in using these data must first submit a research protocol to the Ministry and file a data access request through the *Commission d'accès à l'information*.

Primary and secondary school levels, as well as vocational studies, are governed by the same department. These records notably include any secondary school degree or qualification received, vocational degrees awarded, and the year these degrees were earned. For vocational degrees, the subject is also recorded. From these files, I create an indicator variable for obtaining a secondary school diploma (*DES*) within 5 years of starting secondary school (i.e. the year a student is first observed in grade 7). Note that a student may have been held back in primary school and still obtain a secondary school diploma on time.

The College department records the year a student was first enrolled in any collegial program in Quebec, as well as the program and the institution of that first registration. If a college degree is awarded, the program in which the degree was awarded is recorded (e.g. pre-university degree in Natural Sciences). The exact date the degree was earned is not recorded, however. The files instead indicate whether the degree was completed either (a) on time, (b) less than 2 years after expected duration, or (c) more than 2 years after the expected duration. There is a further caveat: degree completion is only recorded for students who first enrolled in a "normal" college program (*DEC*). For example, degree completion is not recorded for students who first enrolled in a transition program. I use these files to create indicators of college enrollment and college completion. I also approximate the year of completion using the coarse information on time to completion.

The University department records enrollment separately by semester (Fall, Winter and Summer). For each semester, if a student is enrolled in a Quebec university, the number of credits taken, the institution and the field of study are recorded. A separate file is kept for degrees awarded. This file includes the year a degree is awarded, the granting institution, and the type of degree (bachelor, masters, doctoral, 1-year diploma, etc.). With these files, I notably create an indicator of university enrollment and one for bachelor degree completion.

Combining information from all three departments, I then calculate each student's highest level of education. The categories I consider are:

- No secondary school diploma or qualification
- Secondary school diploma (DES)

- Secondary school qualifications
- Vocational degree (DEP)
- Some collegial, started in “normal” program, no degree yet
- Some collegial, did not start in “normal” program
- Pre-university college degree
- Technical college degree
- Other college degree (includes 1-year degrees)
- Some university, no degree yet
- 1-year university diploma
- Bachelor degree or higher

I also calculate each student’s number of years of education. Note that this variable might vary within the categories listed above. For instance, someone who dropped out in grade 9 has 9 years of education, while someone who dropped out in grade 10 has 10. Someone who took 13 years of primary/secondary schooling to obtain a DES and has no further schooling is coded as having 11 years of education (i.e. the normal time it takes to get a DES). Students who were in university for one year and then dropped out have 14 years of education (11 for primary+secondary school, 2 for college, and 1 in university), while those who stayed in university for two years before dropping out have 15 years of education. I top code the number of years of education at 16 (the time it takes to obtain a bachelor degree), however, to avoid my results being driven by outliers. For instance, I do observe a few hundreds students with 19 years of education or more (i.e. people from earlier cohorts in master and PhD programs). The number of years of education therefore incorporates information on multiple margins, e.g. retention in university, college enrollment, vocational studies after secondary school, drop out behavior, etc.

Finally, I create measures of expected earnings. To do so, I calculate earnings percentile ranks (in the national earnings distribution) for all workers aged 30-44 in the Public Use Microdata File of the 2006 Canadian Census , separately by age-group. I then calculate the mean earnings rank for each category of highest level of education, as well as for all possible combinations of level-of-education and field-of-study. Finally, I assign to each student in my data the mean earnings rank asso(Statistics Canada, 2006)ciated with her level of education in the 2006 Canadian Census (or combination of highest level of education and field of study).

Note that students in the 1995 cohort normally finished secondary school in 2005-2006, meaning that 2006 is the year they were making their decision to pursue a post-secondary education.

Measurement of $\Omega^{-i}_{s(i,n)}$ Equation (1) simultaneously includes primary and secondary school fixed effects. This yields one fixed effect for each school in Montreal. Note that students attending a given secondary school need not have attended the same primary school – secondary schools do not nest primary schools.[1]

For each student, I then create a leave-self-out measure for both primary and secondary schools. For instance, for student i and primary school s (which student i attended), I calculate $\Omega^{-i,P}_{s} = \frac{\Omega^{P}_{s(i)}N_s - \tilde{y}_i}{N_s - 1}$, where N_s is the number of permanent residents who attend school s and $\tilde{y}_i = y_i - \bar{y}$ is the deviation of student i's outcome from the sample mean. Student i's outcome must be first re-centered around the sample mean because fixed effects are normalized to have a mean of zero.[2]

Then, I assign the relevant leave-self-out measure to each student-year observation. For years in which a student is in a primary school other than the one he was attending at baseline, no leave-self-out adjustment is necessary since that student was not in that school during the year on which the fixed effect estimation is based. I then take the student-level average of $\Omega^{-i,P}_{s}$ over all primary school years, and similarly calculate a student-level average of $\Omega^{-i,S}_{s}$ for secondary school years. The childhood school quality measure $\Omega^{-i}_{s(i,n)}$ is then the simple sum of these two averages. Note this averaging over primary/secondary school years only matters for permanent residents who have switched school at some point. For the majority of students who only attended one primary and one secondary school, the averaging is redundant, and it is simply the case that $\Omega^{-i}_{s(i,n)} = \Omega^{-i,P}_{s} + \Omega^{-i,S}_{s}$.

In unreported analyses, I use a split-sample approach in which a random half of the sample of permanent residents is used to measure school quality and the other half is used to estimate the regression-discontinuity design. Split-sample and leave-self-out measures of school quality are highly correlated (0.98), hence the results presented in this paper are very similar under the split-sample approach.

Catchment Areas To my knowledge, no electronic, geocoded version of the catchment areas that prevailed in the years 1995-2001 exists. I therefore re-constructed such maps using

[1]Default French primary schools do feed into default secondary schools. But with open enrollment, and the large number of private secondary schools, the connection between local primary and secondary schools is weak.

[2]Jackknife estimates of school fixed effects Ω^{-i}_{s}, in which one regression is ran for each observation, are almost perfectly correlated (0.99) with my hand-calculated leave-self-out measures.

the following procedure.

To first generate a benchmark, the default school associated with each six-digit postal code of the Island of Montreal as of 2015 was recorded by "feeding" each of these ≈45,000 postal codes in the search engines of the websites of the three francophone schools boards. Using shapefiles for Canadian postal codes, I then created a map of all 2015 French catchment areas on the Island of Montreal, down to the six-digit postal code level.

To infer what the boundaries were in the years the cohorts of students I track started grade school, I used two additional sources of information. First, the Ministry of Education provided me temporarily with baseline enrollment data for all 100,929 students in my data set along with their six-digit postal codes (in the analytical data set, six-digit postal codes are de-identified).[3] I then mapped actual attendance patterns and compared with the 2015 boundaries. Second, I used the Internet Archives WaybackMachine (Internet Archive, 2016) to document each school opening/closure that happened since 1995, and extracted old maps of catchment areas from archived versions of the school boards websites (when available).[4] Combining all these sources of information, I deducted where the boundaries must have been drawn, and assigned the appropriate default schools to each postal code by hand. It must be noted that for many schools, the boundaries have not changed since 1995, hence no manual re-coding was necessary. Using ArcGIS, I also calculated, for each postal code, the distance to the nearest boundary and the unique ID of that boundary. Only boundaries that do not coincide with major geographical features, such as highways or canals, were considered. Using these same sources of information, I also inputted catchment areas for English public schools. As explained in the text, however, these boundaries are not well-defined and therefore not used in the analyses.

Attrition About 8% of the total number of students who started grade 1 in Montreal had vanished from primary/secondary school educational records before turning 16. These students are excluded from the main sample used this paper. Interestingly, about 1,000 of these students did enroll in a Quebec university at some point, even though they did not graduate from secondary school in the province. Students who had left Montreal (but remained in Quebec) by the time they turned 15 are also excluded from all analyses.

For higher-education, enrollment in colleges and universities outside the province is not comprised in my dataset. As a result, I may wrongly infer that some students in my main

[3]This first data delivery contained only two variables: school attended (name and code) and postal code of residence. For confidentiality reasons, this file had to be destroyed before the analytical files could be transferred to me.

[4]The five school board websites are www.csdm.ca, www.csmb.qc.ca, www.cspi.qc.ca, www.emsb.qc.ca, and www.lbpsb.qc.ca.

sample never attended college, when in fact they did out-of-province. However, this phenomenon likely only affects a very small proportion of my sample. A few factors provide strong incentives for college and university students to remain in the province, at least for their undergraduate studies. Firstly, tuition fees in Quebec are the lowest in Canada. Secondly, the discrepancies between Quebec's and other North American educational systems generate important timing issues in meeting college requirements. For instance, at the end of secondary school, students in Quebec only have 11 years of schooling, rather than 12. Finally, there is a language barrier that prevents attending school outside of the province for the large vast majority of students who went to primary and secondary school in French.

To assess the possible magnitude of this measurement issue, I use data from the loans and bursaries records of the Ministry of Education. For each year between 1995-1996 and 2014-2015, I was given a series of indicator variables that flag whether student i in my sample was receiving loans or bursaries in year t. Students who resided in Quebec in childhood but go abroad for college are still eligible for loans and bursaries from the Quebec government. Since at the time of enrolling in a foreign college the student's permanent address is often still a Quebec one, it is easier for them to take up loans from Quebec than from another province. I can therefore check the proportion of students who take up students loans while not being enrolled in any postsecondary institution in Quebec to assess the size of the phenomenon. Under this method, I find that about 1% of my sample attended a higher education institution outside the province at some point (many of which also attended a college or a university in Quebec before doing so out-of-province). Finally, it is worth noting that any mis-measurement of educational attainment due to students leaving the province would plausibly lead me to *underestimate* differences across schools and neighborhoods. Students studying abroad, where tuition is much more expensive, are arguably from higher-SES backgrounds, leading me to underestimate educational attainment in places where it is the highest.

Tabulations of enrollments in university and college by province of residents and province of study from Statistics Canada's Postsecondary Student Information System (PSIS) can also help inform the amount of attrition at the post-secondary level.[5] For instance, in 2010, 244,134 students from Quebec were enrolled in university. Of these, 233,634 were in a Quebec institution (96%). The ratio is even higher at the college level.

[5] https://library.queensu.ca/data/educ_tables/psis/

C Mathematical Appendix

C.1 Production Function & Movers Design

The education production function is cumulative and separately additive in yearly family, school, and neighborhood (non-school) inputs:

$$y_i = \sum_{a=0}^{A} \left[\lambda \mu_{n(i,a)} + \omega \psi_{s(i,a)}\right] + A\tilde{\theta}_i$$

where a denotes age, $n(i,a)$ is the location in which child i resided at age a, and $s(i,a)$ the school she attended at that age., and $\tilde{\theta}_i$ are annual average family inputs. Since I am interested in place effects, it is useful to re-write the production function as an exposure-weighted average of inputs received in each location student i ever resided in

$$y_i = \sum_{n} a_{in} \left[\lambda \mu_n + \omega \tilde{\psi}_{s(i,n)}\right] + A\tilde{\theta}_i$$

where a_{in} is the number of years the child resided in location n (with $\sum_n a_{in} = A$), and $\tilde{\psi}_{s(i,n)}$ denotes the annual average quality of schools attended by child i while residing in location n. For instance, for location k, $\tilde{\psi}_{s(i,k)} = \frac{1}{a_{ik}} \sum_{a=0}^{A} \psi_{s(i,a)} 1\{n(i,a) = k\}$.

For one-time movers, educational outcomes are

$$y_i = A\left[\lambda \mu_d + \omega \tilde{\psi}_{s(i,d)} + \tilde{\theta}_i\right] - (m_i - 1) \underbrace{\left[\lambda\left(\mu_d - \mu_o\right) + \omega\left(\tilde{\psi}_{s(i,d)} - \tilde{\psi}_{s(i,o)}\right)\right]}_{e_{i,od}}$$

and the empirical strategy focuses on the second term, which is the average difference in neighborhood (non-school) and school inputs between locations d and o for student i, $e_{i,od}$, interacted with age-at-move m_i. I then re-write $e_{i,od}$ as a function of measurable objects

$$\begin{aligned} e_{i,od} &= \left(\frac{1}{A}\right) \Delta \bar{y}_{od} + \left(\underbrace{\frac{\tilde{\psi}_{s(i,d)} - \tilde{\psi}_{s(i,o)}}{\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR}}}_{c_{i,od}} - 1\right) \omega \left[\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR}\right] - \left[\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}\right] \\ &= \left(\frac{1}{A}\right) \Delta \bar{y}_{od}^{-s} + \left(\frac{c_{i,od}}{A}\right) \pi \Delta \Omega_{od} - \left[\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}\right] + \underbrace{\frac{(c_{i,od} - 1)}{A}\left[A\omega\left(\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR}\right) - \pi \Delta \Omega_{od}\right]}_{\varepsilon_i} \end{aligned}$$

where the term ε_i captures idiosyncratic deviations in changes in true school effects $A\omega\left(\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR}\right)$

from their unbiased forecast $\pi\Delta\Omega_{od}$.

C.2 Forecast-Unbiased School Predicted Gains

To formalize the notion of forecast-unbiasness, it is useful to consider an experimental sample of permanent residents that is randomly assigned to schools and neighborhoods. Their outcomes are given by $y_i^E = A\left[\lambda\mu_n + \omega\tilde{\psi}_{(i,n)} + u_i\right]$, where u_i is uncorrelated with $\mu_{n(i)}$ and $\tilde{\psi}_{s(i,n)}$ by virtue of random assignment. In such a setting, it is possible to estimate the amount of bias in estimates of school effects via a regression of y_i^E on a measure of $\Omega_{s(i,n)}$ constructed using an external, non-experimental sample:

$$y_i^E = \alpha_n + \pi\Omega_{s(i,n)} + \varepsilon_i$$

with

$$\begin{aligned}\pi = \frac{Cov(y_i^E, \dot{\Omega}_{s(i,n)})}{Var(\dot{\Omega}_{s(i,n))})} &= \frac{Cov\left(A\left[\lambda\mu_n + \omega\tilde{\psi}_{s(i,n)} + \tilde{\theta}_i\right], \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})} \\ &= \underbrace{\frac{Cov\left(A\lambda\mu_n, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})}}_{\text{=0 (fixed effects)}} + \frac{Cov\left(A\omega\tilde{\psi}_{s(i,n)}, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})} + \underbrace{\frac{Cov\left(\tilde{\theta}_i, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})}}_{\text{=0 (randomization)}}\end{aligned} \tag{A1}$$

where $\dot{\Omega}_{s(i,n)}$ denotes the residuals from a regression of $\Omega_{s(i,n)}$ on neighborhood fixed effects. In the language of Chetty, Friedman and Rockoff (2014), π is the relationship between true school effects and estimated school effects – it corresponds to the coefficient from the unfeasible regression of $A\omega\tilde{\psi}_{(i,n)}$ on $\Omega_{s(i,n)}$ (conditional on location fixed effects). In comparison, the feasible regression of y_i^{PR} on $\Omega_{s(i,n)}$ as well as a set of neighborhood fixed effects in a non-experimental sample yields a regression coefficient on $\Omega_{s(i,n)}$ of one, by construction. Hence, $1-\pi$ is the amount of forecast bias in estimates of $\Omega_{s(i,n)}$:

$$\frac{Cov\left(y_i^{PR}, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})} = \underbrace{\frac{Cov\left(A\lambda\mu_n, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})}}_{\text{=0 (fixed effects)}} + \underbrace{\frac{Cov\left(A\omega\tilde{\psi}_{s(i,n)}, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})}}_{\pi} + \frac{Cov\left(A\tilde{\theta}_i, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})} = 1$$

Now suppose we did observe μ_n and $\tilde{\psi}_{s(i,n)}$, and ran a regression of y_i^{PR} on both these variables in the *non-experimental* subsample of permanent residents. The regression equation would take the following form:

$$y_i^{PR} = \alpha_n \mu_n + \alpha_s \tilde{\psi}_{s(i,n)} + \epsilon_i. \tag{A2}$$

where ϵ_i is a regression residual and is therefore uncorrelated with school quality by construction. The OLS estimate of α_s is equal to

$$\hat{\alpha}_s = \underbrace{\frac{Cov\left(A\lambda\mu_n, \dot{\tilde{\psi}}_{s(i,n)}\right)}{Var(\dot{\tilde{\psi}}_{s(i,n)})}}_{=0 \text{ (controlling for } \mu_n)} + A\omega + A \underbrace{\frac{Cov\left(\tilde{\theta}_i, \dot{\tilde{\psi}}_{s(i,n)}\right)}{Var(\dot{\tilde{\psi}}_{s(i,n)})}}_{\equiv \rho_s} = A\left(\omega + \rho_s\right)$$

where ρ_s corresponds to the omitted variable bias due to the omission of $\tilde{\theta}_i$ from the regression – ρ_s is a partial regression coefficient in a linear projection of family inputs $\tilde{\theta}_i$ onto μ_n and $\tilde{\psi}_{s(i,n)}$. This suggests the following interpretation: $\pi = \frac{A\omega}{\alpha_s} = \frac{\omega}{\omega+\rho_s}$.[6]

Without an experimental sample, one can still estimate the amount of forecast bias using a valid instrumental variable Z_i that shifts $\tilde{\psi}_{s(i,n)}$ but is otherwise orthogonal to parental inputs $\tilde{\theta}_i$. The IV estimate of the coefficient on $\Omega_{s(i,n)}$ in a regression of y_i^{PR} on $\Omega_{s(i,n)}$ as well as on a set of neighborhood fixed effects is

$$\begin{aligned}\frac{Cov(y_i^{PR}, \dot{Z}_i)}{Cov(\Omega_{s(i,n)}, \dot{Z}_i)} &= \underbrace{\frac{Cov\left(A\lambda\mu_n, \dot{Z}_i\right)}{Cov(\Omega_{s(i,n)}, \dot{Z}_i)}}_{=0 \text{ (fixed effects)}} + \frac{Cov\left(A\omega\tilde{\psi}_{s(i,n)}, \dot{Z}_i\right)}{Cov(\Omega_{s(i,n)}, \dot{Z}_i)} + \underbrace{\frac{Cov\left(A\tilde{\theta}_i, \dot{Z}_i\right)}{Cov(\Omega_{s(i,n)}, \dot{Z}_i)}}_{=0 \text{ (exclusion restriction)}} \\ &= \frac{Cov\left(\pi\Omega_{s(i,n)} + v_i, \dot{Z}_i\right)}{Cov(\Omega_{s(i,n)}, \dot{Z}_i)} = \pi\end{aligned}$$

where v_i is the residual from the unfeasible regression of $A\omega\tilde{\psi}_{s(i,n)}$ on $\Omega_{s(i,n)}$, and $\dot{Z}_i$ denotes the residuals from a regression of Z_i on the neighborhood fixed effects.[7]

[6] Rearranging equation (A2), $\alpha_s \tilde{\psi}_{s(i,n)} = y_i^{PR} - \alpha_n \mu_n - \epsilon_i$, implies that $\frac{Cov\left(\alpha_s \tilde{\psi}_{s(i,n)}, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})} = \frac{Cov\left(y_i^{PR}, \dot{\Omega}_{s(i,n)}\right)}{Var(\dot{\Omega}_{s(i,n)})}$. The LHS is $\frac{\alpha_s \pi}{A\omega}$ and the RHS is 1.

[7] Let $b_s = \pi\Omega_s - A\omega\tilde{\psi}_s$ be the school-level bias of estimated school effects, which is zero on average. Then, $v_i = -b_{s(i,n)}$. The identifying assumption implies that Z_i is only correlated with $\Omega_{s(i,n)}$ through $\tilde{\psi}_{(i,n)}$, i.e. $Cov(b_{s(i,n)}, Z_i) = 0$.

C.3 Horse-Race Decomposition Approach

Consider the following estimating equation

$$y_{icmod} = \beta_{s|n} \left(m_i \times \pi\Delta\Omega_{od}\right) + \beta_{n|s} \left(m_i \times \Delta\bar{y}_{od}^{-s}\right) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod} \quad \text{(A3)}$$

Here, school and non-school *marginal* effects $\beta_{s|n}$ and $\beta_{n|s}$ are identified from variation in the timing of moves. These are partial regression coefficients that reveal the annual effect of a change in one contextual dimension, holding the other constant.[8] Section C.3.2 below provides an interpretation of these reduced-form objects in terms of the parameters of the conceptual model presented in Section IV. The full convergence rate β is a weighted average of these two marginal effects and can be recovered using the following accounting identity:

$$\beta = \beta_{s|n} \left[\frac{Var^r(\pi\Delta\Omega_{od}) + Cov^r(\pi\Delta\Omega_{od}, \Delta\bar{y}_{od}^{-s})}{Var^r(\Delta\bar{y}_{od})}\right] + \beta_{n|s} \left[\frac{Var^r(\Delta\bar{y}_{od}^{-s}) + Cov^r(\pi\Delta\Omega_{od}, \Delta\bar{y}_{od}^{-s})}{Var^r(\Delta\bar{y}_{od})}\right] \quad \text{(A4)}$$

where $Var^r(z)$ and $Cov^r(z)$ respectively denote the variance and covariance of the residuals of $(m_i \times z)$. The total effect of moving to a better area captures independent variation in school and non-school factors (the variances), as well as joint variation in these two dimensions (the covariance). As equation (A4) makes clear, because of possible differences in variances, equal effect size (i.e. $\beta_{s|n} = \beta_{n|s}$) does not imply that schools and other neighborhood factors matter equally. Even if the gains associated with a $\pi\Delta\Omega_{od}$-unit increase in forecast-unbiased school effects are large, schools may nonetheless explain only a small share of the total gains of moving to a better neighborhood if there is little variation in school quality across FSAs (i.e. if $Var^r(\pi\Delta\Omega_{od})$ is small). The empirical counterpart to the school share defined in equation (12) is given by

$$S^{school} = \frac{\beta^{school}}{\beta} = \frac{1}{\beta}\left(\frac{\beta_{s|n} Var^r(\pi\Delta\Omega_{od}) + \beta_{n|s} Cov^r(\pi\Delta\Omega_{od}, \Delta\bar{y}_{od}^{-s})}{Var^r(\Delta\bar{y}_{od})}\right) \quad \text{(A5)}$$

and, correspondingly, the fraction of total gains that is not accounted for by causal school effects is

$$S^{non-school} = \frac{\beta^{non-school}}{\beta} = \frac{1}{\beta}\left(\frac{\beta_{n|s} Var^r(\Delta\bar{y}_{od}^{-s}) + \beta_{s|n} Cov^r(\pi\Delta\Omega_{od}, \Delta\bar{y}_{od}^{-s})}{Var^r(\Delta\bar{y}_{od})}\right). \quad \text{(A6)}$$

For completeness, I also consider an alternative definition of the non-school share which measures the fraction of β that would remain if movers were not to benefit from moving to

[8] Appendix Figure A20 plots estimates of $\beta_{s|n,m}$ and $\beta_{n|s,m}$, which are roughly linear with age-at-move.

places with better schools, which is obtained by setting the marginal effect of geographic differences in school effects on movers, $\beta_{s|n}$, to zero in equation (A4). This approach differs from the main decomposition approach only in how it allocates the covariance terms between the school and non-school components. Results are shown in Table A16. Because the covariance term $Cov^r(\pi\Delta\Omega_{od}, \Delta\bar{y}_{od}^{-s})$ is considerably smaller than the two variances, this alternative approach yields results that are fairly similar to the main approach. In specifications that adjusts for π, the school share range from 53% to 63% for all movers and from 58% to 69% for one-time movers alone.

C.3.1 Derivation of accounting equation (A4)

For ease of exposition, ignore the conditioning variables and fixed effects included in equations eq (4) and (A3), and set $\pi = 1$ so that $(\Delta\bar{y}_{od} - \pi\Delta\Omega_{od}) = \Delta\Lambda_{od}$. Let $m\widehat{\Delta\Omega}_{od}$ denote the residuals of a regression of $m\Delta\Omega_{od}$ on $m\Delta\Lambda_{od}$: $m\widehat{\Delta\Omega}_{od} = m\Delta\Omega_{od} - \frac{Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})}{Var(m\Delta\Lambda_{od})} m\Delta\Lambda_{od}$. Define $m\widehat{\Delta\Lambda}_{od}$ accordingly. Then, the coefficients of the simplified horse-race regression $y_i = \beta_{s|n}(m\Delta\Omega_{od}) + \beta_{n|s}(m\Delta\Lambda_{od}) + \epsilon_i$ are

$$\beta_{s|n} = \frac{Cov\left(m\widehat{\Delta\Omega}_{od}, y_i\right)}{Var(m\widehat{\Delta\Omega}_{od})} \quad ; \quad \beta_{n|s} = \frac{Cov\left(m\widehat{\Delta\Lambda}_{od}, y_i\right)}{Var(m\widehat{\Delta\Lambda}_{od})}.$$

The associated full convergence rate is $\beta = \frac{Cov(y_i, m\Delta\bar{y}_{od})}{Var(m\Delta\bar{y}_{od})}$. Re-organizing

$$
\begin{aligned}
Var(m\Delta\bar{y}_{od}) \times \beta &= Cov\left(m\Delta\bar{y}_{od}, y_i\right) \\
&= Cov\left(m\Delta\Omega_{od}, y_i\right) + Cov\left(m\Delta\Lambda_{od}, y_i\right) \\
&= Cov\left(m\widehat{\Delta\Omega}_{od}, y_i\right) + \frac{Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})}{Var(m\Delta\Lambda_{od})} Cov\left(m\Delta\Lambda_{od}, y_i\right) \\
&+ Cov\left(m\widehat{\Delta\Lambda}_{od}, y_i\right) + \frac{Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})}{Var(m\Delta\Omega_{od})} Cov\left(m\Delta\Omega_{od}, y_i\right) \\
&= \beta_{s|n} Var(m\widehat{\Delta\Omega}_{od}) + \beta_{n|s} Var\left(m\widehat{\Delta\Lambda}_{od}\right) \\
&+ Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od}) \left[\frac{Cov\left(m\Delta\Lambda_{od}, y_i\right)}{Var(m\Delta\Lambda_{od})} + \frac{Cov\left(m\Delta\Omega_{od}, y_i\right)}{Var(m\Delta\Omega_{od})}\right] \\
&= \beta_{s|n} \left[Var(m\Delta\Omega_{od}) - \frac{Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})^2}{Var(m\Delta\Lambda_{od})}\right] \\
&+ \beta_{n|s} \left[Var(m\Delta\Lambda_{od}) - \frac{Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})^2}{Var(m\Delta\Omega_{od})}\right] \\
&+ Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od}) \left[\frac{Cov\left(m\Delta\Lambda_{od}, y_i\right)}{Var(m\Delta\Lambda_{od})} + \frac{Cov\left(m\Delta\Omega_{od}, y_i\right)}{Var(m\Delta\Omega_{od})}\right] \\
&= \beta_{s|n} Var(m\Delta\Omega_{od}) + \beta_{n|s} Var(m\Delta\Lambda_{od}) + \left(\beta_{s|n} + \beta_{n|s}\right) Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})
\end{aligned}
$$

where the non-school share is given by

$$
S^{non-school} = \frac{Cov(y_i, m\Delta\Lambda_{od})}{Cov(y_i, m\Delta\bar{y}_{od})} = \frac{1}{\beta}\left[\frac{\beta_{n|s} Var(m\Delta\Lambda_{od}) + \beta_{s|n} Cov(m\Delta\Omega_{od}, m\Delta\Lambda_{od})}{Var(m\Delta\bar{y}_{od})}\right] = \frac{\beta^{non-school}}{\beta}.
$$

The proof is similar for the more general case in which there are covariates in the estimating equations: First residualize $m\Delta\Omega_{od}$, $m\Delta\Lambda_{od}$ and $m\Delta\bar{y}_{od}$ on the relevant control variables and fixed effects. The accounting identity then becomes

$$
\beta = \frac{1}{Var^r(\Delta\bar{y}_{od})}\left[\beta_{s|n} Var^r(\Delta\Omega_{od}) + \beta_{n|s} Var^r(\Delta\Lambda_{od}) + \left(\beta_{s|n} + \beta_{n|s}\right) Cov^r(\Delta\Omega_{od}, \Delta\Lambda_{od})\right]
$$

where $Var^r(z)$ and $Cov(z)$ denote the variance and covariance of the residuals of $m \times z$.

C.3.2 Interpretation of reduced-form coefficients $\beta_{s|n}$ and $\beta_{n|s}$

For ease of exposition, let the school compliance factor be constant across students: $c_{i,od} = c \forall i$, and let ξ denote the relationship between estimated neighborhood (non-school) effects

Λ_n and true effects $A\lambda\mu_n$. Total exposure effects simplify to

$$\begin{aligned} e_{i,od} =& \frac{1}{A} c\pi\Delta\Omega_{od} + \frac{1}{A}\Delta\bar{y}_{od}^{-s} - \left(\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}\right). \\ =& \frac{1}{A}\left[c - \xi\frac{1-\pi}{\pi}\right]\pi\Delta\Omega_{od} + \frac{1}{A}\left[\xi\right]\Delta\bar{y}_{od}^{-s} \end{aligned} \tag{A7}$$

which implies that $\beta_{s|n} = \frac{1}{A}\left(c - \xi\frac{1-\pi}{\pi}\right)$ and $\beta_{n|s} = \frac{1}{A}\xi$. The first coefficient, $\beta_{s|n}$, is increasing in the compliance rate c, but decreasing in absolute $(1-\pi)$ and relative $\frac{\xi}{\pi}$ sorting into schools. The two coefficients $\beta_{s|n}$ and $\beta_{n|s}$ are equal when $c = 1$ and $\xi = \pi$. As a result, the school share S^{school} is decreasing in the degree of sorting into schools (i.e. it converges to zero as $\rho_s \to \infty$, where ρ_s is defined in section C.2), and the sorting term $\left(\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}\right)$ is allocated to the school and non-school components in proportion of the relative amount of sorting in each dimension. For example, assume full compliance $c = 1$ and let the variance of true school effects be the same as the variance of true neighborhood (non-school) effects: $Var(A\omega\Delta\psi_{od}) = Var(A\lambda\Delta\mu_{od})$. In this case, $\beta_s Var(A\omega\Delta\psi_{od}) > \beta_n Var(A\lambda\Delta\mu_{od})$ if and only if $\pi > \xi$. Similarly, if the degree of sorting is the same in both dimensions $\xi = \pi$ (maintaining the assumption of full compliance), it follows that $\beta_s Var(A\omega\Delta\psi_{od}) > \beta_n Var(A\lambda\Delta\mu_{od})$ if and only if $Var(A\omega\Delta\psi_{od}) > Var(A\lambda\Delta\mu_{od})$. Note that if movers never switch schools $(c = 0)$, β_s could be negative. Intuitively, for given absolute gains due to non-school factors (the numerator of the convergence rate), larger differences in school effects across locations $\pi\Delta\Omega_{od}$ imply a wider gap in educational attainment $\Delta\bar{y}_{od}$ (the denominator of the convergence rate). The convergence rate must therefore be decreasing in the relative importance of school effects if movers don't switch schools.

D School Effects: Specification Checks

D.1 Locally Constant School Effects

Dong and Lewbel (2015) show that in regression discontinuity settings, the change in slope at the cutoff, which they label the treatment effect derivative (TED), has implications for testing whether the LATE is locally constant. In fact, one necessary condition for not having the treatment effect vary with the running variables is that TED=0. In a local linear regression setting, this amounts to testing whether the interaction between the running variable and the treatment dummy is statistically significant. This insight implies that the coefficient on $distance_{ib} \times HighSide_b$ in equations (2) and (3) contains valuable information. In my context, this test indicates whether it is reasonable to assume that school effects are constant with respect to distance to boundaries, and therefore to extrapolate the RD-IV effects away from boundaries.

Table D1: Constant School Effects

Outcome:	University enrollment			DES in 5 years			Years of education		
Boundaries:	All boundaries	Small differences	Large differences	All boundaries	Small differences	Large differences	All boundaries	Small differences	Large differences
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First-Stage ($\Omega^{-i}_{s(n(i))}$)									
RD coefficient (*HighSide*)	0.0328	0.0159	0.0470	0.0337	0.0132	0.0539	0.1511	0.0916	0.2340
	(0.0065)	(0.0089)	(0.0092)	(0.0061)	(0.0077)	(0.0101)	(0.0298)	(0.0401)	(0.0454)
distance	0.0247	0.0171	0.0303	0.0326	0.0188	0.0480	0.1209	0.1379	0.0976
	(0.0118)	(0.0151)	(0.0175)	(0.0097)	(0.0096)	(0.0201)	(0.0517)	(0.0564)	(0.1065)
distance X HighSide	-0.0133	-0.0080	-0.0081	-0.0282	-0.0200	-0.0098	-0.0791	-0.0897	-0.1011
	(0.0176)	(0.0211)	(0.0292)	(0.0128)	(0.0110)	(0.0342)	(0.0705)	(0.0757)	(0.1885)
Reduced-form									
RD coefficient (*HighSide*)	0.0279	0.0131	0.0422	0.0347	0.0165	0.0528	0.1165	0.0652	0.1830
	(0.0087)	(0.0117)	(0.0129)	(0.0084)	(0.0106)	(0.0142)	(0.0390)	(0.0497)	(0.0643)
distance	0.0293	0.0207	0.0387	0.0328	0.0193	0.0481	0.1315	0.2094	0.0585
	(0.0162)	(0.0212)	(0.0257)	(0.0158)	(0.0155)	(0.0342)	(0.0715)	(0.0738)	(0.1458)
distance X HighSide	-0.0040	0.0090	-0.0254	-0.0279	-0.0200	-0.0178	-0.0759	-0.1389	-0.0817
	(0.0221)	(0.0303)	(0.0368)	(0.0176)	(0.0176)	(0.0455)	(0.0805)	(0.0812)	(0.2322)
RD-IV									
RD coefficient (*HighSide*)	0.8542	0.8242	0.9017	1.0340	1.2509	0.9837	0.7739	0.7107	0.7857
	(0.1645)	(0.4683)	(0.1740)	(0.1618)	(0.5492)	(0.1801)	(0.1575)	(0.3257)	(0.1733)
distance	0.0082	0.0067	0.0114	-0.0010	-0.0043	0.0008	0.0374	0.1109	-0.0189
	(0.0101)	(0.0146)	(0.0153)	(0.0094)	(0.0121)	(0.0188)	(0.0437)	(0.0569)	(0.0761)
distance X HighSide	0.0073	0.0155	-0.0179	0.0014	0.0051	-0.0080	-0.0139	-0.0747	-0.0005
	(0.0122)	(0.0182)	(0.0202)	(0.0098)	(0.0123)	(0.0266)	(0.0481)	(0.0511)	(0.1094)
N	43291	22928	20362	43291	22928	20362	43291	22928	20362
Cohort fixed effects	x	x	x	x	x	x	x	x	x
Individual characteristics	x	x	x	x	x	x	x	x	x
Neighborhood (FSA) fixed effects	x	x	x	x	x	x	x	x	x
Boundary fixed effects	x	x	x	x	x	x	x	x	x

Notes: This table replicates Table 2, but using alternative bandwidth restrictions and subsamples. To examine the generalizability of baseline estimates of causal effect of school quality on educational attainment away from the cutoff, in columns (1), (4) and (7) the sample includes all permanent residents with no bandwidth restriction. In the remaining columns, the sample is split in half according to the size of the difference in school quality $\hat{\Omega}^P_{s(i)}$ between the two default options on opposing sides of a given boundary. In columns (2), (5) and (8), only boundaries with differences below the mean are in the sample. In columns (3), (6) and (9), the sample is restricted to above-average boundaries in terms of the magnitude of the differences between the two default French primary schools. All standard errors are clustered at the French primary school boundary level. See notes to Table 3 for additional details

In Table D1, coefficients on $distance_{ib}$ and $distance_{ib} \times HighSide_b$ from a RD regression that includes all permanent residents (since we are interested in extrapolating to the full sample) are reported in columns (1), (4) and (7) for university enrollment, DES in 5 years and years of education, respectively. For convenience, distance is measured in kilometers (rather than meters). The reduced-form results indicate that educational attainment is increasing with distance, plausibly because (a) households further away from boundaries live closer to the school itself, and (b) households at larger distances are generally located in suburban areas where educational attainment is relatively higher.

The coefficient on the interaction term $distance_{ib} \times HighSide_b$ is generally negative but smaller in magnitude than the main effect of distance, indicating that outcomes increase with distance on both sides of boundaries. However, both for the first-stage and the reduced-form, these interactions are small and not statistically significant. The interactions in the RD-IV second-stage equation are also small and insignificant, consistent with locally constant effects.

D.2 Linearity of π

In this section, I examine whether the assumption of a linear relationship between estimated and true school effects is reasonable. To do so, I split the sample of boundaries used in the RD analysis in two according to the size of the difference in quality between the two default French primary schools. In columns (2), (5) and (8) of Table D1, the RD-IV model is estimated on the subsample of students who live near small-gap boundaries. Similarly, in columns (3), (6) and (9), the sample is restricted to boundaries with above-average differences in quality between the two default options.

For all three outcomes, large-gap boundaries are associated with an average jump in childhood school quality $\Omega^{-i}_{s(i,n)}$ more than twice the size of the jump around small-gap boundaries. For example, the first-stage coefficient for university enrollment is 0.016 for small-gap boundaries, and 0.047 for large-gap boundaries. Similarly, the reduced-form RD coefficients for large-gap boundaries are more than twice the size of the coefficients for small-gap boundaries. As a result, the estimates of π are fairly constant across "gap size". For university enrollment, while the main RD-IV estimate is 0.85, it is only slightly smaller for small-gap boundaries (0.82) and slightly higher for large-gap boundaries (0.90). For DES in 5 years, the estimate of π is surprisingly large for small-gap boundaries (1.25), but this coefficient is not very precisely estimated (s.e. 0.55). For years of educations, π also appears to be stable across the two sets of boundaries (0.71 and 0.79).

D.3 Model-based approximation of π

As in section C.3.2, let the school compliance factor c_i be constant across students, and let ξ denote the relationship between estimated neighborhood (non-school) effects Λ_n and true effects $A\lambda\mu_n$. In this case, we can write

$$e_i(o,d) = \left(\frac{c\pi}{A}\right)\Delta\Omega_{od} + \left(\frac{\xi}{A}\right)\Delta\Lambda_{od}. \tag{A8}$$

Now suppose we run the horse-race model in equation (A3), but use $\Delta\Omega_{od}$ and $\Delta\Lambda_{od}$ as regressors. Under the above assumptions, the partial regression coefficient on $\Delta\Omega_{od}$ is equal to $\frac{c\pi}{A}$. Provided one knows the values of c and A, it is therefore possible to approximate the value of π using the movers design alone. There is one important caveat: in practice, the compliance rate is certainly not constant and likely interacts with $\Delta\Omega_{od}$ (e.g. moves on short distances might be associated with small differences in school quality as well as with low probabilities of switching school) as well as with age-at-move m_i. Hence, the *estimated* partial regression coefficient at best approximates $\frac{c\pi}{A}$.

To obtain a rough estimate of c, I conduct an event-study similar to the one described in Section VI.B. For movers, I define an index of relative school quality by

$$\sigma^{\psi}_{od(i,t)} = \frac{\Omega_{s(i,t)} - \overline{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}} \tag{A9}$$

where $\Omega_{s(i,t)}$ is the quality of the school attended by student i at time t (measured by the fixed effects estimates obtained in Section I.D), and $\overline{\Omega}_{s(n,t)}$ is the average quality of schools attended by permanent residents of FSA n at time t. The corresponding event-study results are shown in Figure A21. The index increases sharply in value right at the time of the move. While there seems to be a modest spike in the year preceding the move, this bump is very small compared to the break that occurs on impact. Pre-post differences in the index of relative school quality for one-time movers therefore provide a rough estimate of the compliance rate.[9] The estimates reported in Figure A21 for one-time movers range between

[9] Let the average index of relative school quality $\sigma^{\psi}_{od(i,t)} = \frac{\Omega_{s(i,t)} - \overline{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}$ for pre-move years, that is for years spent in the origin, be $\frac{\tilde{\Omega}_{s(o,t)} - \overline{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}$. Similarly, the average for post-move years is $\frac{\tilde{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}$, and the difference between the two is $\frac{\tilde{\Omega}_{s(d,t)} - \tilde{\Omega}_{s(o,t)}}{\overline{\Omega}_{s(d,t)} - \overline{\Omega}_{s(o,t)}}$. Note that the plots that include all movers may suffer from measurement error because the index of relative school quality is likely mismeasured in pre-move years for people who move multiple time before reaching their final destination.

0.54 and 0.65. I therefore calibrate $c = 0.6$ below.

Model-based estimates of π are shown in Table D2. In columns (1) and (3), I calibrate the exposure period to $A = 10$ (grade 1 to grade 10), and in columns (2) and (4) I set $A = 12$ to encompass all school years, that is including kindergarten and the last year of secondary school.

Using reasonable values of A and c, I obtain estimates of π that range between 0.72 and 0.87 for university enrollment, between 0.62 and 1 for timely secondary school graduation, and between 0.72 and 0.89 for years of education. These values are quite similar to the ones obtained using the RD design, which also range between 0.6 and 1 across specifications and outcomes.

Table D2: Movers-based Estimates of π

Sample:	All movers		One-time movers	
	(1)	(2)	(3)	(4)
Coefficient on (m_i × $\Delta\Omega_{od}$)				
University enrollment	-0.0431	-0.0431	-0.0435	-0.0435
	(0.0097)	(0.0097)	(0.0128)	(0.0128)
Secondary school diploma in 5 years	-0.0374	-0.0374	-0.0502	-0.0502
	(0.0102)	(0.0102)	(0.0138)	(0.0138)
Years of education	-0.0445	-0.0445	-0.0432	-0.0432
	(0.0106)	(0.0106)	(0.0128)	(0.0128)
Calibration				
A	10	12	10	12
c	0.6	0.6	0.6	0.6
Implied value of π				
University enrollment	0.72	0.86	0.73	0.87
Secondary school diploma in 5 years	0.62	0.75	0.84	1.00
Years of schooling	0.74	0.89	0.72	0.86
Cohort fixed effects	x	x	x	x
Individual characteristics	x	x	x	x
Age at move fixed effects	x	x	x	x
Origin-by-destination fixed effects	x	x	x	x
N	24316	24316	15533	15533

Notes: This table presents estimates of the causal effect of school quality on educational attainment (π) based on the conceptual model in Section IV. I first report estimates of the partial effect of exposure to a location with higher educational attainment schools obtained using the movers design. In columns (1) and (2), the partial regression coefficients on $m_i \times \Delta\Omega_{od}$, denoted by $\tilde{\beta}_s$, are estimated for all movers, and in columns (3) and (4) they are estimated on the subsample of one-time movers. The implied value of π is calculated by $\pi = -\tilde{\beta}_s A/c$. I assume a value of the compliance rate c of 0.6 given the evidence shown in Figure A21. The number of exposure years A is calibrated to 10 in columns (1) and (3), and to 12 in columns (2) and (4).

D.4 House price analysis

Unfortunately, I cannot look at house prices at a fine level of geography because the the 6-digit postal code variable is anonymized in the analytical dataset. As a coarse alternative, I assign average house prices to each student in the data on the basis of the census tract in which they reside at baseline. To facilitate the comparison of magnitudes with the previous literature, measures of childhood school quality are standardized to have mean zero and variance of one. In column (1), the sample includes all permanent residents and the reported coefficients show the raw association between school quality $\Omega^{-i}_{s(i,n)}$ and house prices. A one s.d increase in school quality is associated with a 0.11 increase in log house prices. Restricting the sample to students within the baseline RD bandwidths used in the paper and controlling for neighborhood fixed effects brings the coefficient down to about 0.02, but it remains strongly statistically significant (column 2). Next, I estimate the RD model. Column (3) is the first-stage on standardized school quality, which shows that students on the better side of a boundary do attend better schools on average. But in column (4), I find no evidence that they reside in tracts with greater house prices. If anything, the coefficients are negative, but not statistically significantly so.

Table D3: House Price Analysis

	Dependent variable:			
	ln(House price)	ln(House price)	School quality (std)	ln(House price)
	(1)	(2)	(3)	(4)
University enrollment				
School quality ($\Omega^{-i}_{s(i,n)}$), standardized	0.1105	0.0224		
	(0.0149)	(0.0046)		
RD coefficient			0.1110	-0.0225
			(0.0311)	(0.0130)
Secondary school diploma in 5 years				
School quality ($\Omega^{-i}_{s(i,n)}$), standardized	0.0985	0.0208		
	(0.0112)	(0.0034)		
RD coefficient			0.1103	-0.0099
			(0.0282)	(0.0130)
Years of education				
School quality ($\Omega^{-i}_{s(i,n)}$), standardized	0.1131	0.0194		
	(0.0141)	(0.0042)		
RD coefficient			0.1007	-0.0116
			(0.0320)	(0.0126)
Sample restriction:	All PRs	RD BW	RD BW	RD BW
Cohort fixed effects		x	x	x
Individual characteristics			x	x
Neighborhood (FSA) fixed effects		x	x	x
Boundary fixed effects			x	x

Notes: This table examines the relationship between school quality and house prices. House prices are measured at the census tract level and obtained from the 2001 Canadian Census. In columns (1) and (2), ln house prices are regressed on the average quality of schools attended during childhood $\Omega^{-i}_{s(i,n)}$ (standardized with mean zero and variance of one in the sample of permanent residents). The is no control variable in column (1), whereas cohort and FSA fixed effects are added in column (2). Columns (3) and (4) replicate the regression discontinuity design of Table 2. In column (3), childhood school quality is the outcome, whereas ln house prices are in column (4). All standard errors are clustered at the French primary school boundary level. See notes to Table 2 for additional details on the regression discontinuity approach.

E Level of Geography

Here, I verify that the results do not hinge on the choice of level of geography used to define neighborhoods. To do so, I consider census tracts (CT) as an alternative unit. There are about 500 different census tracts in Montreal, more than five times the number of FSAs.

By construction, there is more raw variation in educational attainment across CTs than across FSAs. However, when estimating school quality using a two-way fixed effect model (equation (1)), the amount of variation across schools, net of neighborhood fixed effects, is largely unaffected by the choice of geography. As shown in Table (E1), even when one increases the number of neighborhood units five-fold, it is still the case that there is more variation across schools than across neighborhoods.

Table E2 then presents estimates of total exposure effects when neighborhoods are defined as census tracts. In Panel A, origin-by-destination fixed effects are included. The number of effective observations in the first two columns (18,981) is considerably smaller than for the results based on FSA (24,316 movers) because there are multiple cases of origin-by-destination pairs that, at the census tract level, contain only one observation, and are therefore dropped. I therefore consider a less restrictive specification in Panel B, where origin and destination fixed effects enter separately.

In columns (1) and (2) all movers are included. The reported convergence rates tend to vary between 2% and 2.5%, considerably smaller than at the FSA level. This is likely due to the fact that sampling error is greater at smaller levels of geography (i.e. fewer permanent residents per neighborhoods), and perhaps also reflect greater sorting of permanent residents at smaller levels of geography. Census tracts may also less precisely capture all features of the community in which children live and socialize. Chetty et al. (2018) also obtain a convergence rate of 2.5% when using census tracts as the main unit of analysis.

In columns (3) and (4), I only consider movers who, at least once, moved between FSAs, and in column (5) and (6) I only consider children who moved across census tracts but remained within the same FSA. For between-FSA moves (but still measuring neighborhood quality at the census tract level), I now find convergence rates around 2.5-3%, slightly higher than rates for all movers, but still considerably smaller than FSA-level estimates. For within-FSA moves, I find no evidence of any convergence – no estimate is statistically significant at conventional levels, and in Panel A coefficients even turn positive. For completeness, in columns (7) and (8) I restrict the sample to one-time movers, and only focus on census tract movers who moved across FSAs exactly once in columns (9) and (10). Convergence rates for these sub-groups vary between 2% and 3%.

Finally, in Table E3, I investigate whether the school share is smaller or larger when using census tracts rather than FSAs. The total exposure effects I decompose are those shown in Panel B, column (1) of Table E2. Note that for consistency, the entire procedure is re-estimated at the census tract level. In other words, the school fixed effects used in this analysis are net of census tract fixed effects, and than the RD-IV coefficient of π is similarly estimated substituting census tract fixed effects for FSA fixed effects in equations (2) and (3) .

Interestingly, in column (1), the school share S^{school} is slightly smaller at the census tract level than at the FSA level for university enrollment (70%) and secondary school graduation (67%), but the opposite is true for years of education (86%). The estimates that are adjusted for sorting into schools, presented in column (3), are respectively 53%, 76% and 73% for university enrollment, secondary school graduation and years of education. Overall, the estimates are neither systematically larger nor smaller than at the FSA level, but rather similar.

Two opposing forces push the school share in opposite directions when going from the FSA to the census tract level. First, because there are 5 times more census tracts than FSAs, the variance of non-school factors $Var^r(\Delta\bar{y}_{od}^{-s})$ becomes relatively larger than the variance of school factors $Var^r(\pi\Delta\Omega_{od})$, which tends to make the school share smaller. However, because there might be more sorting of permanent residents across census tracts than across FSAs, and because census tract may only provide a noisy estimate of true neighborhood quality, β_n will likely be much smaller at the census tract level than at the FSA level (relative to β_s), making the non-school share smaller. I find that this is indeed the case. At the FSA level, it is generally the case that $\beta_n > \beta_s$, but the opposite is generally true at the census tract level.

Table E1: Variation Across Census Tracts and Schools

	Outcome					
	University enrollment		DES in 5 years		Years of education	
	(1)	(2)	(3)	(4)	(5)	(6)
Student-level standard deviation of fixed effects:						
Neighborhoods (Census Tracts)	0.159	0.081	0.152	0.068	0.734	0.328
Schools	0.247	0.235	0.261	0.255	1.170	1.123
Dependent variable summary statistics:						
Mean	0.460		0.729		13.323	
Standard deviation	[0.498]		[0.444]		[2.083]	
Fixed effects estimated						
Separately	x		x		x	
Simultaneously		x		x		x
Number of students	37,491					
Number of primary schools	435					
Number of secondary schools	211					
Number of neighborhoods	502					

Notes: This table replicates Table 1, defining neighborhoods as census tracts rather than FSAs. The sample is restricted to students who always resided in the same census tract (permanent residents). See notes to Table 1 for additional details.

Table E2: Total Exposure Effects: Moves Across Census Tracts

Sample:	All movers		Moved across FSAs at least once		Never moved across FSAs		One-time movers		Moved across FSAs exactly once	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Measure of educational attainment										
					Panel A: Origin-by-destination fixed effects					
University enrollment	-0.0182	-0.0218	-0.0252	-0.0290	0.0093	0.0084	-0.0171	-0.0225	-0.0153	-0.0206
	(0.0116)	(0.0114)	(0.0137)	(0.0132)	(0.0251)	(0.0255)	(0.0191)	(0.0189)	(0.0190)	(0.0186)
Secondary school diploma in 5 years	-0.0175	-0.0200	-0.0196	-0.0222	0.0001	0.0019	-0.0182	-0.0201	-0.0300	-0.0355
	(0.0123)	(0.0115)	(0.0141)	(0.0133)	(0.0237)	(0.0228)	(0.0190)	(0.0182)	(0.0202)	(0.0196)
Years of education	-0.0173	-0.0219	-0.0261	-0.0315	0.0142	0.0132	-0.0101	-0.0147	-0.0194	-0.0268
	(0.0115)	(0.0112)	(0.0136)	(0.0130)	(0.0246)	(0.0247)	(0.0172)	(0.0167)	(0.0186)	(0.0180)
N	18981	18981	12442	12442	5618	5618	9062	9062	7460	7460
					Panel B: Origin + destination fixed effects					
University enrollment	-0.0225	-0.0236	-0.0250	-0.0259	-0.0152	-0.0171	-0.0287	-0.0315	-0.0281	-0.0313
	(0.0064)	(0.0063)	(0.0067)	(0.0065)	(0.0194)	(0.0194)	(0.0091)	(0.0090)	(0.0082)	(0.0080)
Secondary school diploma in 5 years	-0.0222	-0.0224	-0.0253	-0.0251	-0.0086	-0.0095	-0.0168	-0.0189	-0.0366	-0.0388
	(0.0066)	(0.0061)	(0.0067)	(0.0062)	(0.0196)	(0.0189)	(0.0096)	(0.0092)	(0.0084)	(0.0080)
Years of education	-0.0272	-0.0277	-0.0305	-0.0309	-0.0102	-0.0109	-0.0266	-0.0298	-0.0306	-0.0341
	(0.0061)	(0.0059)	(0.0064)	(0.0061)	(0.0197)	(0.0194)	(0.0082)	(0.0082)	(0.0079)	(0.0076)
N	31333	31333	24777	24777	6522	6522	15921	15921	15469	15469
Cohort fixed effects	x	x	x	x	x	x	x	x	x	x
Individual characteristics	x	x	x	x	x	x	x	x	x	x
Age at move fixed effects	x	x	x	x	x	x	x	x	x	x
Only moved once					x	x	x	x	x	x
Times in difficulty before moving		x		x		x		x		x

Notes: This table replicates Table 2, estimating total exposure effects from movers across census tracts. Coefficients shown in the table are convergence rates β. Standard errors are clustered at the destination census tract level. Columns (1) and (2) include all students moving across census tracts within Montreal. In columns (3) and (4), the sample is restricted to census tract movers who moved across a FSA boundary at least once, and in columns (5) and (6) it is restricted to those who never moved across FSAs. Columns (7) and (8) restrict the sample to student who moved only once, and in columns (9) and (10) the sample is restricted to those who moved across FSAs exactly once (but may have moved across census tracts within FSAs multiple times). See notes to Table 2 for additional details on the movers design.

Table E3: Decomposition: Census Tract Level Analysis

Sample:	All movers		
	(1)	(2)	(3)
University enrollment			
	Total exposure effects		
β	-0.0225	-0.0225	-0.0225
	(0.0064)	(0.0064)	(0.0064)
	School and non-school components		
β^{school}	-0.0158	-0.0161	-0.0099
	(0.0048)	(0.0048)	(0.0030)
$\beta^{non\text{-}school}$	-0.0067	-0.0064	-0.0126
	(0.0040)	(0.0041)	(0.0046)
Share school effects (S^{school})	70%	72%	44%
	(0.1394)	(0.1418)	(0.0871)
Secondary school diploma in 5 years			
	Total exposure effects		
β	-0.0222	-0.0222	-0.0222
	(0.0066)	(0.0066)	(0.0066)
	School and non-school components		
β^{school}	-0.0149	-0.0140	-0.0149
	(0.0052)	(0.0052)	(0.0055)
$\beta^{non\text{-}school}$	-0.0073	-0.0083	-0.0073
	(0.0038)	(0.0038)	(0.0038)
Share school effects (S^{school})	67%	63%	67%
	(0.1348)	(0.1354)	(0.1447)
Years of education			
	Total exposure effects		
β	-0.0272	-0.0272	-0.0272
	(0.0061)	(0.0061)	(0.0061)
	School and non-school components		
β^{school}	-0.0233	-0.0230	-0.0161
	(0.0052)	(0.0051)	(0.0036)
$\beta^{non\text{-}school}$	-0.0039	-0.0041	-0.0110
	(0.0033)	(0.0035)	(0.0038)
Share school effects (S^{school})	86%	85%	59%
	(0.1087)	(0.1121)	(0.0784)
Measure of school quality	$\pi\Omega_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$	$\pi\Omega^{-i}_{s(n)}$
π	1	1	RD estimate

Notes: This table replicates Table 4, but using census tracts to define neighborhoods. Sample restricted to movers across census tracts within Montreal. Standard errors are clustered at the destination census tract level, and obtained by the Delta method. Parameters β^{school}, $\beta^{non-school}$ and S^{school} are calculated using the methods presented in section V. In columns (1) and (2), π is set to one. In column (3), π is given by RD-IV estimates obtained using the methods described in the notes to Table 2 but for the sample of census tract permanent residents. In column (1), the difference in average school quality between permanent residents of areas d and o is given by $\Delta\Omega_{od} = \bar{\Omega}_d^{PR} - \bar{\Omega}_o^{PR}$, while in remaining columns $\Delta\Omega_{od}$ is measured using census tract-level averages of permanent residents' leave-self-out childhood school quality $\bar{\Omega}_n^{-i}$.

References

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica*, 82(6): 2295–2326.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *The American Economic Review*, 106(8): 2110–2144.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review*, 104(9): 2593–2632.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." National Bureau of Economic Research 25147.

Dong, Yingying, and Arthur Lewbel. 2015. "Identifying the effect of changing the policy threshold in regression discontinuity models." *Review of Economics and Statistics*, 97(5): 1081–1092.

Internet Archive. 2016. "Wayback Machine." (accessed September 1, 2016).

Jann, Ben. 2017. "KMATCH: Stata module for multivariate-distance and propensity-score matching." *Statistical Software Components, Boston College Department of Economics.*

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics*, 142(2): 698–714.

Statistics Canada. 2006. "Census of Canada, 2006, Individuals File." Statistics Canada [producer]. Using CHASS [distributor].

Steven Manson, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2019. "IPUMS National Historical Geographic Information System: Version 14.0 [Database]." (accessed October 15, 2019).