

Long-Term Contextual Effects in Education: Schools and Neighborhoods[†]

By JEAN-WILLIAM LALIBERTÉ*

This paper estimates the long-term impact of growing up in better neighborhoods and attending better schools on educational attainment. First, I use a spatial regression-discontinuity design to estimate school effects. Second, I study students who move across neighborhoods in Montreal during childhood to estimate the causal effect of growing up in a better area (total exposure effects). I find large effects for both dimensions. Combining both research designs in a decomposition framework, and under key assumptions, I estimate that 50–70 percent of the benefits of moving to a better area on educational attainment are due to access to better schools. (JEL H75, I21, R23)

Educational outcomes vary greatly across regions, neighborhoods and schools. Given the sizable economic and nonpecuniary benefits to education, disparities in educational attainment can translate into persistent socioeconomic inequality in adulthood.¹ What is the relative importance of different channels in driving long-term education outcomes? Multiple policy interventions target neighborhoods directly or incentivize families to relocate to low-poverty areas, motivated by the belief that social context significantly influences students' aspirations and learning. Schools are key institutions of local communities, plausibly constituting a pivotal mechanism fueling spatial inequalities.

*University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4 (email: jeanwilliam.lalibert@ucalgary.ca). John Friedman was coeditor for this article. I am very grateful to Natalie Bau, Kory Kroft and Philip Oreopoulos for their guidance throughout this project. Thanks to Josh Angrist, Gustavo Bobonis, Raj Chetty, David Deming, Nicolas Gendron-Carrier, Michael Gilraine, Hugh Macartney, Richard Mansfield, Ismael Mourifié, Matthew Notowidigdo, Mathieu Marcoux, Juan Morales, Scott Orr, Rob Oxoby, Uros Petronijevic, Marc-Antoine Schmidt, Michel Serafinelli, Aloysius Siow, and Alex Whalley for insightful discussions and comments. I also thank Simon Bézy, Latifa Elfassihi, and Sophie LeBoutillier from the Ministère de l'Éducation et de l'Enseignement supérieur for their assistance with the data. Financial support from the Social Sciences and Humanities Research Council (SSHRC) grant 752-2016-1150, and the Ontario Graduate Scholarship (OGS) is gratefully acknowledged. All mistakes are my own.

[†]Go to <https://doi.org/10.1257/pol.20190257> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

¹On the economic returns to education see Oreopoulos and Petronijevic (2013); Moretti (2004); Card (2001); and Angrist and Krueger (1991); and, specifically for Canada, Boudarbat, Lemieux, and Riddell (2010). These returns are particularly large for marginal students (Zimmerman 2014). On nonpecuniary benefits, see Heckman, Humphries, and Veramendi (2017); Oreopoulos and Salvanes (2011); and Milligan, Moretti, and Oreopoulos (2004).

Empirical evidence on the relative importance of schools and of neighborhoods for educational attainment remains scarce, despite the important implications of such information for the allocation of public resources toward policies directed at either schools or neighborhoods. This is due to both a lack of data linking students to long-term outcomes and also several key econometric challenges faced by those who compile such data. For instance, in many jurisdictions, school attendance is strictly residence-based, making these two dimensions observationally indistinguishable. Estimation of neighborhood and school effects is further complicated by sorting of families.

This paper evaluates the long-term impact of growing up in better neighborhoods and attending better schools. To do so, it combines student-level administrative data with several key institutional features of Quebec's education system to overcome the data and identification challenges that have hindered analyses of school and neighborhood effects. The large longitudinal database follows students who grew up in the region of Montreal throughout their entire educational career and tracks them on a yearly basis as they switch schools, move across neighborhoods, and make higher education investments.

I undertake three sets of analyses in this paper. First, I estimate the causal effect of school quality on long-term educational outcomes using a boundary-based regression discontinuity design. Second, I use a movers design to estimate *total exposure effects*—the combined effect of an additional year of exposure to a given neighborhood and to its schools. In both analyses, I find sizable contextual effects. Third, I develop an empirical framework that, under several strong assumptions on effect heterogeneity, allows me to combine the two research designs to decompose the total gains in educational attainment of growing up in a given area into school and non-school neighborhood components. The results indicate that total exposure effects are large but mostly driven by schools rather than by other non-school neighborhood characteristics.

Quebec is particularly well suited for investigating the role of schools independently of place of residence. The province operates two parallel public school systems—one French and one English—thereby allocating neighbors to different default neighborhood schools on the basis of their mother tongue. Crucially, the catchment area boundaries for French and English schools are different, allowing separate identification of schools from other neighborhood effects.

Empirically, I first construct measures of school *predicted gains* (school quality) representing school-level differences in educational attainment that cannot be accounted for by where students reside. Then, I use local variation in the quality of French school default options to recover forecast-unbiased school effects by instrumenting for school quality in a spatial regression-discontinuity framework (RD-IV), leveraging the fineness of the spatial information in the administrative files. I then check the validity of the RD-IV approach by showing that the educational outcomes of students attending English schools are smoothly distributed around *French* primary school boundaries, thereby confirming that these boundaries do not coincide with discontinuous changes in non-school unobserved attributes.

Next, I use students who move during childhood to estimate the magnitude of total exposure effects. To address the endogeneity of location decisions, I exploit

variation in the *timing* of within-city moves across families.² Intuitively, if exposure matters, the educational outcomes of movers should converge toward those of the permanent residents of their destination (children who always resided in the same area) with increasing time spent in that location. The reduced-form object of interest is a convergence rate. To insure the estimates are not confounded by sorting of movers into different areas, the model relies on comparisons between children who started in the same neighborhood and moved to the same neighborhood, but did so at different ages. The main empirical specification includes both origin-by-destination and age-at-move fixed effects, with the identifying assumption that the degree of selection into locations does not vary systematically with children's age at the time of the move. To provide support for this assumption, I conduct a series of robustness checks, notably family fixed effects specifications and controlling for time-varying observables around the time of the move.

My estimates suggest that movers' educational attainment converges linearly at an annual rate of about 4.5 percent toward the outcomes of the permanent residents in their destination neighborhood. Put differently, moving one year earlier to a place where permanent residents have one more year of education than those of one's origin location increases own educational attainment by 0.045 years. Extrapolating over 10 years of compulsory schooling, these effects account for almost half (45 percent) of the differences in outcomes of permanent residents between origin and destination. The magnitude of these effects is remarkably similar to that reported in Chetty and Hendren (2018a) and Chetty et al. (2018) despite important differences between our two settings.

Finally, to decompose total exposure effects into school and non-school components, I partition the mean outcomes of permanent residents into a part reflecting the average quality of local schools and a non-school residual.³ With forecast-unbiased measures of school quality in hand, one can predict by how much movers' outcomes would improve on the basis of spatial differences in school quality alone, and evaluate the share of total gains of moving that can be accounted for by these school effects. I find that between 50 percent and 70 percent of the effect on educational attainment of moving to a given area are due to access to better schools. In other words, once we take out differences in causal school effects between origin and destination, the rate of convergence on the remaining differences associated with non-school factors is less than half the size of the total convergence rate of 4.5 percent. This suggests that if average school quality was equalized across all neighborhoods, the gains associated with growing up in a relatively higher educational attainment area would be more than halved.

² Aaronson (1998) and Weinhardt (2014) also use variation from movers to identify neighborhood effects. Chetty and Hendren (2018a, b) use rich tax data to track families with children who move across commuting zones and counties in the United States and estimate the causal effect of places on earnings. Similar identification strategies have also been used to analyze health care utilization (Finkelstein, Gentzkow, and Williams 2016), physician practice style (Molitor 2018), the impact of EITC on labor supply (Chetty, Friedman, and Saez 2013), and brand preferences (Bronnenberg, Dubé, and Gentzkow 2012).

³ This empirical approach is consistent with a simple education production function in which cumulative school and non-school neighborhood inputs are additively separable and treatment effects are constant across students. Analyses presented in Sections ID and VI suggest these functional form assumptions are plausible in my setting. Section IV describes the corresponding conceptual framework in details.

The validity of this decomposition hinges on several strong assumptions. First, RD estimates of school effects need to be representative of the average school effects experienced by movers. I implement several tests of constant school effects in Section VI that support the generalizability of these estimates. Second, I assume that neighborhood and school effects are the same for movers and permanent residents. One possible violation of that assumption would be contextual effects that vary by income level; since permanent residents and movers likely exhibit systematic income differences, movers design estimates might suffer from attenuation bias, thereby only corresponding to a lower bound on place effects. Intuitively, if exposure effects vary across movers and permanent residents, projecting permanent residents' outcomes onto movers will not accurately capture the causal effects for movers. While I cannot account for parental income differences, in Section IV, I provide some suggestive evidence that exposure effects are on average similar for movers and permanent residents. For example, the correlation in neighborhood-level average outcomes between the two groups is very close to one. A final, more general caveat to these results is that I can only examine the relative importance of schools and other neighborhood factors for educational attainment; these results could differ for other long-term outcomes—such as earnings or incarceration—even for the same population in Montreal.

This paper brings together several literatures. First, it speaks directly to the classic question: “Do neighborhoods matter?” Correlational analyses generally find strong associations between neighborhood poverty and success at school (Sharkey and Faber 2014, Burdick-Will et al. 2011). In contrast, most experimental and quasi-experimental studies that tackle the challenging task of isolating place effects from nonrandom sorting of families into neighborhoods have found limited evidence of static neighborhood effects on educational and economic outcomes (Ludwig et al. 2013; Kling, Liebman, and Katz 2007; Oreopoulos 2008, 2003; Jacob 2004).⁴ In a recent reanalysis of the Moving to Opportunity (MTO) experiment, Chetty, Hendren, and Katz (2016) show that children do benefit from moving to better locations both in terms of earnings and college enrollment, but that these gains only materialize for youth who moved before the age of 13, consistent with cumulative exposure effects. Similarly, Chetty and Hendren (2018a,b) estimate large exposure effects for children moving across US commuting zones and counties. Given that school attendance is generally residence-based, these estimates of neighborhood exposure effects also reflect differences in local school quality (Altonji and Mansfield 2014). My estimates of total exposure effects are consistent with this prior literature, and my decomposition results suggest neighborhood exposure effects on educational attainment operate mostly via schools.

Second, my paper relates to a parallel stream of research evaluating the causal impact of schools on educational and labor market outcomes. Large effects of attending better schools are found using quasi-experiments (Gould, Lavy, and Paserman

⁴Notable exceptions include Goux and Maurin (2007), who find positive effects of neighborhood peers on the probability of repeating a grade using variation from public housing projects in France; Gould, Lavy, and Paserman (2011), who find significant effects of childhood conditions on adult outcomes in Israel; and Damm and Dustmann (2014), who find that exposure to neighborhood crime in childhood increases convictions later in life in Denmark.

2004), lottery-based designs (Angrist et al. 2017, Deming et al. 2014, Dobbie and Fryer 2015, 2011), and admission threshold rules (Pop-Eleches and Urquiola 2013, Jackson 2010). Using similar research designs, however, Abdulkadiroğlu, Angrist, and Pathak (2014) and Cullen, Jacob, and Levitt (2006), respectively, find no positive effects of attending an elite school or of attending one's preferred school in a school choice program. My paper takes a different approach, exploiting spatial discontinuities in the spirit of Black (1999), to show the early schooling environment has a long-term impact: residing on the better side of a French primary school boundary at age six affects educational outcomes measured more than ten years later.

I also contribute to a growing body of research contrasting the magnitude of school and neighborhood effects.⁵ Historically, researchers have generally focused on either schools or communities, with a few review papers speculating on the relative effectiveness of school and neighborhood interventions by comparing separate studies. Fryer and Katz (2013) and Katz (2015) contrast results from the MTO experiment (Ludwig et al. 2013; Kling, Liebman, and Katz 2007), which induced low-income families to move to low-poverty neighborhoods, with the effects of the Harlem Children's Zone experiment, which combines both school-level and community-level interventions (Dobbie and Fryer 2015, 2011). They conclude that school interventions are likely more effective than community programs for educational outcomes, a conclusion also reached by Oreopoulos (2012). At larger levels of aggregation, Rothstein (2017) finds differences in quality of K–12 education (measured by test scores) account for little of between-city differences in intergenerational mobility, while Card, Domnisoru, and Taylor (2018) find that state- and county-level school quality was a key factor driving regional differences in upward mobility in the early twentieth century. My paper adds to this evidence by showing that school quality goes a long way explaining why place matters for educational attainment.

The methods used in this paper have several empirical benefits. Movers likely constitute a more diverse cross-section of the population than samples of experimental studies that focus on very disadvantaged households (Oreopoulos 2003) or negatively selected populations of lottery applicants (Chyn 2018), contributing to the external validity of the results. Also, by using outcome-based measures of neighborhood and school quality, I circumvent the issue of choosing which observable characteristics to use to proxy for quality. For instance, school input measures and teacher observable characteristics often fail to predict effectiveness, despite the evidence that both schools and teachers have large causal effects on student outcomes (Dobbie and Fryer 2013; Chetty, Friedman, and Rockoff 2014; Rivkin, Hanushek, and Kain 2005; Hanushek 1986).

⁵In sociology, Carlson and Cowen (2015) examine short-run variation in test score gains across schools and neighborhoods in Milwaukee, and Wodtke and Parbst (2017) explore how school poverty mediates neighborhood effects on math and reading tests in the PSID. Sykes and Musterd (2011) study how school characteristics mediate the relationship between neighborhood characteristics and test scores in the Netherlands. In economics, Card and Rothstein (2007) separately examine the effects of school and neighborhood segregation on test scores, and Billings, Deming, and Ross (2016) consider the role of school and neighborhood peers in the formation of criminal networks.

The rest of the paper proceeds as follows. First, I describe the institutional context and the data in Section I. Estimates of causal school effects are presented in Section II, and the movers design used to estimate neighborhood exposure effects is presented in Section III. I set up a conceptual framework that lays the foundations for the decomposition exercise in Section IV, and empirically implement the decomposition in Section V. A host of robustness checks are conducted in Section VI, and Section VII concludes.

I. Data and Background

A. Quebec's Education System and Social Context

Levels of Education.—In Quebec, education is compulsory from age 6 to 16, and most children enroll in kindergarten at age 5. Children complete six years of primary education (grades 1–6), and then attend a secondary school for five more years (grades 7–11), until obtaining a secondary school diploma (*diplôme d'études secondaires, DES*), or equivalent qualifications. Grade repetition is common and over 20 percent of students drop out of secondary school before obtaining any degree.

The higher education system differs considerably from standard North American systems.⁶ In Quebec, there is a sharp hierarchical distinction between *college* and *university*, the former being a prerequisite for the latter. After secondary school, most students enroll in college in either a pre-university (two years) or technical program (three years). The typical student obtaining a pre-university college degree then enrolls in a three-year bachelor degree program in university. In the empirical application, I measure neighborhood and school exposure up until the academic year a student is aged 15 on September 30, inclusive. All educational investments made after that point are considered outcomes.

School Choice between Sectors.—Quebec's education system possesses multiple elements of school choice that contribute to breaking the mechanical link between area of residence and school attendance. At the primary and secondary levels, two public school systems operate in parallel—one French and one English. Public schools are governed by school boards, which are responsible for personnel, transportation, infrastructure, and the allocation of resources across schools. School boards are language-specific, with every residential address falling within the territory of one English and one French school board.⁷ Importantly, the attendance zones of English and French schools are not the same. Hence, two neighbors with different mother tongues who both attend their nearest language-specific public school likely have school peers who originate from different neighborhoods. Access to instruction in English is restricted to anglophones born in Canada, a

⁶Online Appendix Figure A1 shows the typical education course toward a bachelor's degree in Quebec and in a standard North American system. No transition between levels of education in Quebec coincide with the age at which students transition in other educational systems. The number of years of education associated with a bachelor degree, however, remains the same.

⁷Before 1998, school boards were religion-specific (Catholic or Protestant), but individual schools were still either French or English.

strictly enforced rule where parents must obtain an eligibility certificate before enrolling a child in an English school. In the language of the law, anglophones are students whose mother or father attended an English primary or secondary school in Canada. Almost all immigrants are de facto forbidden from attending English schools. Exceptions to the rule are rare.⁸

In contrast with other Canadian provinces and the United States, private schools are widespread in Quebec, notably at the secondary school level. In Montreal, almost a third of all students attend a private secondary school. Private schools do not have attendance zones and are relatively accessible given generous subsidies. Subsidized private schools are also subject to the language of instruction restriction.⁹

School Choice within Sectors.—Quebec's open enrollment policy stipulates that parents have the right to enroll their child in the school of their choice (*libre choix*), subject to capacity constraints and language restrictions. In practice, school boards assign children to default neighborhood schools, and parents who desire to enroll their child in a public school other than the one they are assigned must complete the relevant paperwork at the neighborhood school. Default options may induce two sets of parents living in the same area to enroll their children in different schools since catchment area boundaries often cut through neighborhoods. In this paper, I focus exclusively on French primary school boundaries since English primary school boundaries are not as well defined—some English schools offer different programs (e.g., English Core versus Bilingual) and their catchment areas often vary by program. French primary school boundaries serve as the basis for a regression-discontinuity analysis described in Section II.¹⁰

Importantly, over the time period studied here, there existed no public information about relative primary school quality and performance, such as rankings on outcome-based measures. All Montreal school boards strongly oppose public disclosure of rankings or quality indicators at the primary school level.¹¹ If enrollment exceeds capacity, priority is given to children residing in the school's catchment area and to siblings of children attending the school, and students opting out of their assigned school are not eligible for school bus transportation. Other nonresidence based admission criteria are used for elite magnet schools. The neighborhood school therefore acts as a default option, and catchment area boundaries as cost shifters. In

⁸Language restrictions do not apply to post-secondary institutions.

⁹Nonsubsidized English schools are allowed to enroll non-English speaking students. However, these schools are uncommon and represent less than 1 percent of total enrollment (Duhaime-Ross 2015). Among subsidized private schools, very few charge the maximum fee allowed by law (Lefebvre, Merrigan, and Verstraete 2011). A minority of these schools have entrance exams, yet the vast majority of students taking such exams are admitted to their preferred school (Lapierre, Lefebvre, and Merrigan 2016).

¹⁰Online Appendix Figure A2 shows that French primary school boundaries often cut through census tracts. At the secondary school level, English public schools in Montreal do not have catchment areas, but French public schools do.

¹¹Secondary school rankings are published yearly in mainstream media. I therefore focus exclusively on primary school boundaries for identification. In 2015, a well-known newspaper published partial rankings of Montreal public primary schools for the first time. The cohorts of students analyzed in this paper had left primary school many years before that.

my data, every neighborhood school enrolls at least some students residing outside its catchment area.¹²

Comparison between Montreal and other North American Contexts.—How does Montreal compare to large North American cities? Online Appendix Table A3 lists mean census tract characteristics for Greater Montreal, one similarly sized city (Boston), and one large city with open enrollment in public schools (Chicago), and the United States more broadly. On average, Montreal has a lower employment rate, lower educational attainment and slightly more single-headed families than Boston and Chicago. Montreal exhibits approximately the same amount of between-tract variation in these measures as Boston does, while there is more segregation across tracts in Chicago. The key difference is that Chicago, and Boston to a lesser degree, exhibit considerably more racial segregation than Montreal, where spatial differences are instead explained by segregation by income and immigration status. Montreal also differs from large US cities in that crime rates are substantially lower (Gannon 2001). To the extent that crime and racial segregation might be important drivers of neighborhood effects in the United States, caution is warranted in generalizing the results in this paper to the United States.

B. Data

The main data consist of student-level administrative records provided by Quebec's Ministry of Education (Ministère de l'Éducation, 2016) covering all levels of education (primary school to university). Separate files from four different branches of the Ministry were matched using unique student identifiers. For each year students are enrolled in primary and secondary education, school attended (both public and private), grade level, and the six-digit postal code of residence are recorded. Postal codes (very small geographic areas, generally equivalent to a block-face or a unique apartment building) determine the default neighborhood schools (one English and one French). Catchment areas were manually geocoded on that basis.

In addition to the assigned neighborhood schools, I calculated for each postal code the distance to the nearest catchment area boundary, distance to the nearest public school, associated census tract and Forward Sortation Area (FSA; postal code-based neighborhoods constituting the main geographic unit of analysis).¹³ All distances were calculated separately for the English and French public school systems. In Montreal, students reside in over 500 different census tracts and about 100 FSAs. For confidentiality reasons, school identifiers and six-digit postal codes are

¹² One important reason why capacity constraints do not appear to be binding is that Quebec's school system was experiencing a decline in school-age population over the time period covered here, which notably led to several public school closures in the early 2000s.

¹³ An FSA is defined by the first three digits of a postal code. FSAs are regularly used to operationalize neighborhoods in Canadian research (Card, Dooley, and Payne 2010) and are sometimes used for determining eligibility to community programs. For example, the *Pathways to Education* program targeting residents of the neighborhood of Pointe-Saint-Charles is available to households living in one specific FSA. In my data, educational attainment is relatively smoothly distributed spatially within FSAs: only 3 percent of the within-FSA residual variation occurs between census tracts. Online Appendix E examines the robustness of exposure effects estimates to using census tracts instead of FSAs to define neighborhoods.

de-identified in the analytical dataset. Student demographics—age on September 30, gender, mother tongue, country of origin, language spoken at home—are included, in addition to time-varying variables such as school day care use (primary school only) and an indicator of whether a student is currently considered to have learning difficulties (primary and secondary school).¹⁴ In addition, I append census tract-level characteristics from the 2001 Canadian Census.

In terms of long-term educational outcomes, the data include enrollment and graduation information for secondary school, and all vocational, college, and university programs. I use these to calculate—among other outcomes—university enrollment, timely secondary school graduation (*DES* in five years), and number of years of education. More detailed information regarding the construction of the outcome variables is provided in the Data Appendix.

The sample is focused on residents of the Island of Montreal, Quebec's most populous region and main urban center. This territory encompasses three francophone and two anglophone school boards, and includes the city of Montreal and a few smaller municipalities located in the suburban westernmost part of the Island or enclaved within the city of Montreal. These municipal divisions are irrelevant for school resources administration purposes.

Administrative records were obtained for five cohorts of students who started primary school between 1995 and 2001, following students until the 2014–2015 academic year.¹⁵ The sample consists of all students who resided on the Island of Montreal when entering grade 1 (100,929 students). This selection rule ensures that every student's entire education history is known, and therefore excludes students who moved to Montreal after completing grade 1 elsewhere. On average, about 210 students per FSA enter grade 1 on any given year. The main sample (92,764 students) excludes all students who left Quebec's education system before turning 16.¹⁶

C. Descriptive and Summary Statistics

Descriptive statistics by mobility status are shown in Table A1. Permanent residents (PR) are defined as those who, by the age of 15, had always resided in the same FSA. I distinguish between movers who were still living in Montreal by age 15 and those who had moved off the Island but remained in the province. Because of the within-city focus of this paper, students who left Montreal are excluded from the empirical analyses.

¹⁴ On any given year in primary and secondary school, students with social maladjustment or learning disabilities can be identified as being "in difficulty." School boards receive extra funding to support these students, and many observers worry that schools may "overdiagnose" students as a result. Yet, the predictive power of this variable with respect to educational attainment is stunning. The probability that one obtains a secondary school diploma on time decreases monotonically with each year flagged in difficulty (Figure A3). For the two earliest cohorts, the probability of obtaining a bachelor's degree is 36 percent for students never identified in difficulty, while it is only 5 percent among those who were flagged at least once.

¹⁵ Data for the 1997 and 1999 cohorts are not available.

¹⁶ In primary and secondary school, attrition is generally due to students leaving the province. In exceptional cases, some students may disappear from administrative records if they attend illegal schools, or are home schooled. Students leaving the system are disproportionately non-French speakers and immigrants. Several analyses presented in the Data Appendix indicate that attrition is unlikely to affect my results.

In Montreal, students are on average 6 years old when they enter primary school. Only half the sample consists of native French speakers, but 75 percent of students attend school in French. *Allophones*—defined as individuals whose mother tongue is neither French nor English—make up almost a third of the sample. Nevertheless, the vast majority of students are born in Canada (90 percent). Anglophones are over-represented among permanent residents, while francophones are disproportionately more likely to move outside of Montreal, and allophones to move within Montreal. At baseline (in grade 1), 4 percent of students are considered to have learning difficulties (flagged “in difficulty”), and the fraction increases sharply over time. By the time they reach the age of 15, almost a third of the sample will have been flagged at least once. In general, movers appear to be negatively selected: In grade 1, 3 percent of permanent residents are in difficulty, while 5 percent of movers are.

The number of years for which I track students varies across cohorts, hence observed educational attainment is higher for earlier cohorts, by construction. Online Appendix Table A2 reports summary statistics for some educational outcomes separately by cohort. Roughly 76 percent of students obtain a secondary school diploma (*DES*), but only 61 percent do so on time (in five years), with little variation across cohorts. The college enrollment rate is consistent across cohorts, at 70 percent.¹⁷ In terms of university-level outcomes, as of 2015, 46 percent of students who started primary school in 1995 had enrolled in university and 28 percent had completed a bachelor degree. Virtually no student of the 2001 cohort has a bachelor degree yet, but 22 percent of them are enrolled in university. Every econometric model in this paper includes cohort fixed effects to account for these differences.

Given the variety of school choice options available in Montreal, students living in the same neighborhood need not attend the same school. For instance, at baseline, students living in the average FSA attend as many as 57 different primary schools.¹⁸ When entering grade 1, 63 percent and 50 percent of students in French and English schools attend their neighborhood school, respectively. In total, 41 percent of students opt out of their default option at baseline, with this proportion exceeding 70 percent by the end of secondary school.¹⁹ Opt out rates vary between the primary and secondary school levels primarily because of differences in availability of private school options. Around 12 percent of Montreal students are in the private sector in primary school, a fraction rising to almost 30 percent in secondary school (Figure A6). Yet, geography remains an important factor for many parents when it comes to deciding which school their child will attend. For example, among students in French schools at baseline, 68 percent attend their default school if that school is the nearest French public school from their house, while only 50 percent do so if it is not.²⁰ Despite the open enrollment policy, there is a large discontinuity in school attendance around boundaries at baseline, suggesting that many parents passively

¹⁷ For reference, the college enrollment rate is also 70 percent in Chetty and Hendren’s (2018a) US-wide sample.

¹⁸ See online Appendix Figure A5 for distribution of FSAs by number of schools.

¹⁹ By definition, students in English secondary schools all opt out since English public schools do not have attendance zones at the secondary level. At the age of 15, 58 percent of students in French schools attend a school other than their default option.

²⁰ For 30 percent of students in French schools, the default option is not the nearest French public primary school.

select the default option. This is consistent with a body of evidence in behavioral economics and psychology on the importance of default options (Chetty 2015; Lavecchia, Liu, and Oreopoulos 2014).²¹

D. Measurement

Here, I explain how school and neighborhood quality is measured, and use these measures to describe the amount of variation across schools and neighborhoods in the data. Formally, I estimate a two-way fixed effects model on the subsample of permanent residents:

$$(1) \quad y_i = \Omega_{s(i)} + \Lambda_n + \delta_c + \varepsilon_i,$$

where y_i is a long-term educational outcome for student i from cohort c , living in neighborhood n , and attending the set of schools $s(i)$. The model includes cohort (δ_c), FSA (Λ_n), and school ($\Omega_{s(i)}$) fixed effects. Intuitively, this model is identified because the set of students living in the same area attend a variety of different schools, and students in the same school reside in different neighborhoods.²² Since students generally attend two different schools during childhood—one primary and one secondary school—I parameterize the vector of school effects to include a fixed effect for primary school attended at baseline (Ω_s^P) and a fixed effect for secondary school attended at age 15 (Ω_s^S). I therefore obtain a proxy for school quality for each school in the dataset. Note that these measures of school quality are net of neighborhood fixed effects and therefore reflect the contribution of schools (and sorting into schools) that cannot be accounted for by where schools gather their students from. Primary school quality is net of the secondary schools its students will eventually attend, and secondary school quality is net of the primary schools it gathers its students from. These outcome-based measures of school quality can be interpreted as predicted gains and reflect any observed and unobserved differences in productive school inputs—e.g., teacher and principal quality. Traditional measures of school quality based on test scores may not fully capture other important dimensions of school effectiveness for long-term educational attainment, such as effects on noncognitive skills (Jackson 2016; Heckman, Stixrud, and Urzua 2006).

Table 1 describes the amount of variation in the data for the three main outcomes of interest: university enrollment, finishing secondary school on time (*DES in 5 years*), and years of education. In columns 1, 3, and 5, I first report the raw standard deviation of mean neighborhood-level and school-level outcomes. Total SD for neighborhood effects is the standard deviation of simple (cohort-adjusted) mean

²¹ Discontinuity in school attendance at French primary schools boundaries is shown in online Appendix Figure A4. To create the figure, I randomly pick one of the two schools that share a boundary, separately for each boundary. I then calculate the fraction of students who attend the randomly chosen school as a function of distance to the nearest boundary. Students at positive distances are residing in the catchment area of the randomly chosen school. On the left side of the border (negative distances), 20 percent of students attend the school located on the other side rather than their own default option or any other French school.

²² Just like models of worker and firm fixed effects are identified from switchers (Abowd, Kramarz, and Margolis 1999), this model requires that students of a given neighborhood be observed in multiple schools and that students from a given school be observed in multiple neighborhoods.

TABLE 1—VARIANCE DECOMPOSITION FOR SCHOOL AND NEIGHBORHOOD FIXED EFFECTS

	Outcome					
	University enrollment		DES in five years		Years of education	
	(1)	(2)	(3)	(4)	(5)	(6)
Neighborhood fixed effects						
Total standard deviation	0.139	0.062	0.138	0.046	0.680	0.258
Noise standard deviation	0.042	0.038	0.041	0.032	0.183	0.148
Signal standard deviation	0.133	0.049	0.132	0.033	0.655	0.212
Reliability	0.910	0.624	0.910	0.517	0.928	0.672
School fixed effects						
Total standard deviation	0.249	0.235	0.270	0.264	1.207	1.141
Noise standard deviation	0.058	0.066	0.056	0.062	0.242	0.269
Signal standard deviation	0.242	0.226	0.264	0.257	1.182	1.108
Reliability	0.946	0.921	0.957	0.945	0.960	0.944
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 estimates of variance components of educational attainment of students who always resided in the same neighborhood (permanent residents). To produce these estimates, I regress measures of educational attainment on school, neighborhood, and cohort fixed effects, and report the student-level standard deviation of those fixed effects. Columns 1, 3, and 5 school and neighborhood fixed effects are estimated in separate regressions. In columns 2, 4, and 6, all fixed effects are estimated simultaneously from equation (1). School fixed effects are measured by the sum of a primary and a secondary school fixed effect. The total standard deviation is the student-level standard deviation of school or neighborhood fixed effects. The noise standard deviation is the square-root of the average squared standard error, where standard errors are obtained using 300 bootstrap replications, clustering at the school-by-neighborhood level, and taking the standard deviation of these 300 sets of fixed effects estimates. The signal standard deviation is the square root of (total variance – noise variance). The reliability is the ratio of the signal variance to total variance. In columns 1 and 2 the outcome is university enrollment, in columns 3 and 4 it is an indicator for completing secondary school in five years (*DES* in five years), and in columns 5 and 6 it is years of education.

outcome of PRs, \bar{y}_n^{PR} . Educational attainment varies dramatically across neighborhoods of Montreal: a one standard deviation increase in neighborhood-level university enrollment rates is equal to 14 percentage points, relative to a sample mean of 44 percent. Figure 1 maps these differences. The gap between neighborhoods with best and worst outcomes is abysmal, with fractions of students enrolling in university ranging from 15 percent to 80 percent.²³

For all three outcomes, the SD across schools is almost twice as large as the SD across FSAs.²⁴ These estimates are all relatively precise, with less than 9 percent

²³ Even starker disparities emerge across census tracts (online Appendix Figure A7).
²⁴ This result is not due to the fact that there are fewer FSAs than schools, as the patterns replicate at the census tract level (online Appendix Table E1). Also, these patterns closely reflect the conclusions of Carlson and Cowen (2015), who focus on the variance in test scores growth across neighborhoods and schools in Milwaukee’s open enrollment system. Agrawal, Altonji, and Mansfield (2018) also find that schools account for a larger share of the variance than neighborhoods.

Panel A. University enrollment

Panel B. DES in five years

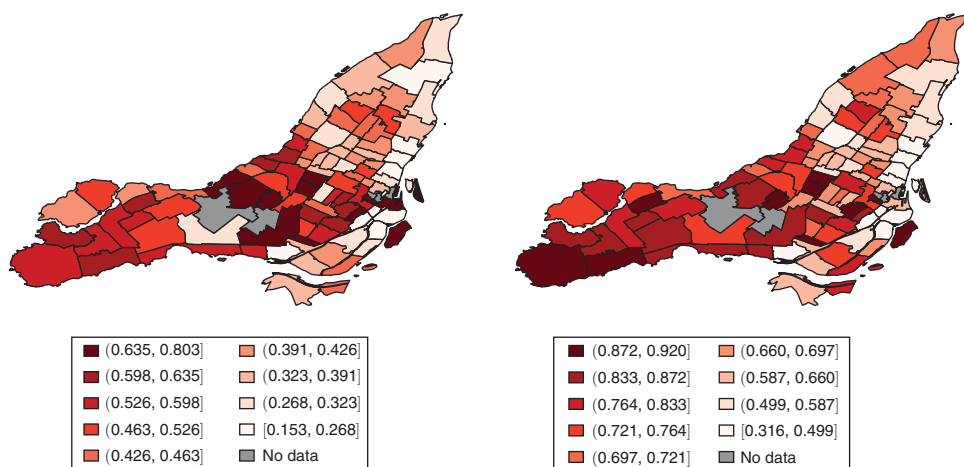


FIGURE 1. SPATIAL VARIATION IN EDUCATIONAL OUTCOMES

Notes: These maps plot FSA-level average educational attainment for students who always resided in the same Forward Sortation Area (permanent residents). Mean outcomes are neighborhood (FSA) fixed effects from a regression of educational attainment on FSA and cohort fixed effects. The fixed effects are then re-centered so that their average is equal to the unconditional sample mean. FSAs with fewer than 10 permanent residents are in gray (labeled “No data”). University enrollment is equal to one for students who were ever enrolled in a Quebec university, and zero otherwise. DES in five years is equal to one for students who completed secondary school in five years, and zero otherwise.

(5 percent) of the variation in neighborhood-level (school-level) mean outcomes due to sampling error. In columns 2, 4, and 6, I report the standard deviation of neighborhood and school fixed effects estimated from the two-way specification (equation (1)). While the magnitude of the variation across schools barely changes when FSA fixed effects are included, a large fraction—between 55 and 65 percent—of the raw standard deviation across FSAs is accounted for by school attendance. It is worth noting that estimates of $\Omega_{s(i)}$ are much more precisely estimated (reliability of 0.92–0.94) than estimates of Λ_n (reliability of 0.52–0.67). I return to the issue of sampling error in Section VIC.²⁵

Two additional stylized facts are worthy of mention. Firstly, the student-level correlation between school and FSA fixed effects is positive but small (e.g., 0.17 for years of education).²⁶ Online Appendix Figure A8 shows the spatial variation in FSA-level averages of $\hat{\Omega}_{s(i)}$ and $\hat{\Lambda}_n$. The two maps exhibit little overlap—places with low average school quality do not necessarily have a low non-school

²⁵ By construction, $\bar{y}_n^{PR} = \bar{\Omega}_n^{PR} + \Lambda_n$, where $\bar{\Omega}_n^{PR}$ is the average of $\Omega_{s(i)}$ for permanent residents of neighborhood n . To maintain this mapping intact in the decomposition analyses, I work with unadjusted estimates of school and neighborhood fixed effects. Section VIC extends the main empirical results to split-sample IV and empirical Bayes shrinking techniques to adjust for sampling error.

²⁶ The correlations are 0.05 and 0.13 for graduating secondary school on time and university enrollment, respectively. The correlations are slightly higher if one uses empirical Bayes shrunk estimates to account for sampling error: 0.20, 0.08, and 0.15 for years of education, timely secondary school graduation, and university enrollment, respectively.

quality component $\hat{\Lambda}_n$. Secondly, there is little evidence of systematic interactions between the two contexts. Following the approach developed in Card, Heining, and Kline (2013), Figure A9 is constructed by slicing the distributions of school and FSA fixed effects into deciles, and then plotting the average residuals in each school-by-neighborhood decile cell. Most average residuals are smaller than 0.1 year of schooling, or less than 5 percent of a standard deviation in the sample of permanent residents. If there were positive interactions between school and neighborhood quality, one would expect abnormally large and positive mean residuals for cells corresponding with high or low deciles in both dimensions. The figure shows no such discernible pattern, which lends support to the additive separability specification. In addition, allowing for unrestricted match effects between schools and neighborhoods (i.e., a full set of indicator variables for each possible combination of neighborhood and primary/secondary school) only slightly improves the model's fit—e.g., for years of education, the *adjusted* R^2 for equation (1) increases from 0.3710 to 0.3735.

The descriptive evidence suggests there is independent variation across both schools and neighborhoods that cannot be accounted for by the other dimension. The variance components reported in Table 1 may partly reflect sorting into schools and neighborhoods. My empirical objectives are to estimate the fraction of these variances that are causal.

II. Effect of Attending Better Schools

Estimation Framework.—In this section, I use a RD-IV approach to estimate the causal effect of attending high predicted gains schools on long-term educational outcomes. Identification is based on the fact that schools' catchment areas cut through neighborhoods in such ways that students on opposite sides of a shared boundary reside in the same community and enjoy the same neighborhood amenities (Black 1999; Billings, Deming, and Rockoff 2014). These boundaries shift the quality of schools two neighbors may be exposed to by varying their default option. I focus on the nearest French *primary* school boundary from one's residence at baseline throughout.²⁷

For each boundary, I first identify which of the two default schools is of greater quality, as measured by fixed effect estimates of predicted gains $\hat{\Omega}_s^P$ in educational attainment. Note that because these fixed effects are net of FSA-level variation and secondary school attendance, the “better” school for a given boundary is not necessarily the one where students have the best outcomes in absolute terms. For each student, I define an indicator variable $HighSide_{ib}$ for whether student i resides on the better side of the nearest French primary school boundary b .²⁸ These indicator

²⁷ Given the importance of private schools, there are considerably fewer boundaries than schools.

²⁸ For over a quarter of all permanent residents the boundary-specific higher quality default school is not the one with relatively higher absolute outcomes. In other words, if I were to assign values of $HighSide_{ib}$ on the basis of absolute outcomes rather than of adjusted school quality $\hat{\Omega}_s^P$, the values of the dummy would flip for a fourth of my sample. Absence of sorting at boundaries on the basis of *adjusted* school quality is consistent with findings that parental preferences are unrelated to school effectiveness once peer quality is accounted for (Abdulkadiroğlu et al. 2017, Rothstein 2006). Similarly, while school test scores are capitalized in house prices in residence-based school

variables are used as instruments in the following two-stage regression-discontinuity framework:

$$(2) \quad y_{icnb} = \pi \Omega_{s(i,n)}^{-i} + f(\text{distance}_{ib}) + \gamma X_{icnb} + \alpha_b + \alpha_n + \alpha_c + \varepsilon_{icnb},$$

$$(3) \quad \Omega_{s(i,n)}^{-i} = \zeta \text{HighSide}_b + f(\text{distance}_{ib}) + \gamma X_{icnb} + \alpha_b + \alpha_n + \alpha_c + \varepsilon_{icnb},$$

where (3) and (2) are first- and second-stage equations. The dependent variable y_{icnb} is an educational outcome for permanent resident i of neighborhood n . Student-level individual characteristics X_{icnb} are included to improve precision—the point estimates are virtually identical if these covariates are omitted. Each student is matched to the boundary b that is the nearest from her home. The main regressor of interest, $\Omega_{s(i,n)}^{-i}$, is a leave-self-out measure of cumulative school quality over i 's entire childhood.²⁹ The coefficient π represents the average causal effect of a one unit increase in school quality.

In both stages, a control function for distance to the nearest boundary $f(\text{distance}_{ib})$ is included, as well as FSA (α_n), boundary (α_b), and cohort (α_c) fixed effects. In the main specification, I follow Lee and Lemieux (2010) and parameterize $f(\text{distance}_{ib})$ with a rectangular kernel. Standard errors are clustered at the French primary school boundary level. The main results use optimal bandwidths based on the procedure developed in Calonico, Cattaneo, and Titiunik (2014). I examine the robustness of the results to functional form assumptions and bandwidth restrictions in Section VIA.

The validity of the RD approach rests on the assumption that right around boundaries, the quality of default school options is as good as random. In education systems where school attendance is fully determined by residence, households may sort right around boundaries, generating discontinuities in sociodemographic characteristics (Bayer, Ferreira, and McMillan 2007). However, in Montreal, opportunities to opt out of one's default public school and the availability of private schools strongly reduce any incentive to sort at boundaries. For instance, Fack and Grenet (2010) find that the capitalization of school quality in house prices in Paris falls sharply with private school availability, and is effectively zero in areas with many private schools. More importantly, given that rankings of Montreal primary schools are not publicly available, distinguishing good from bad nearby schools is difficult and parents may have little ability to sort at boundaries.³⁰

To validate that any jump in school quality at boundaries does not reflect discrete changes in student characteristics, I verify that observable characteristics are balanced around these boundaries. Combining all covariates to generate measures of

attendance systems, school *value-added* is often not (Imberman and Lovenheim 2016; Kane, Riegg, and Staiger 2006), with some exceptions (Gibbons, Machin, and Silva 2013).

²⁹The childhood school quality measure $\Omega_{s(i,n)}^{-i}$ is obtained by taking the sum of leave-self-out transformations of the primary (Ω_s^P) and secondary school (Ω_s^S) fixed effects estimated in Section ID. The exact procedure is described in the online Data Appendix. Jackknife and split-sample approaches yield almost identical results.

³⁰Online Appendix Figure A10 shows a density plot by distance to boundaries. No excess density is observed on the higher quality side of the threshold (side with schools with greater predicted gains on university enrollment). A formal McCrary (2008) test finds no statistically significant gap: the log difference in height is 0.006 with a standard error of 0.018.

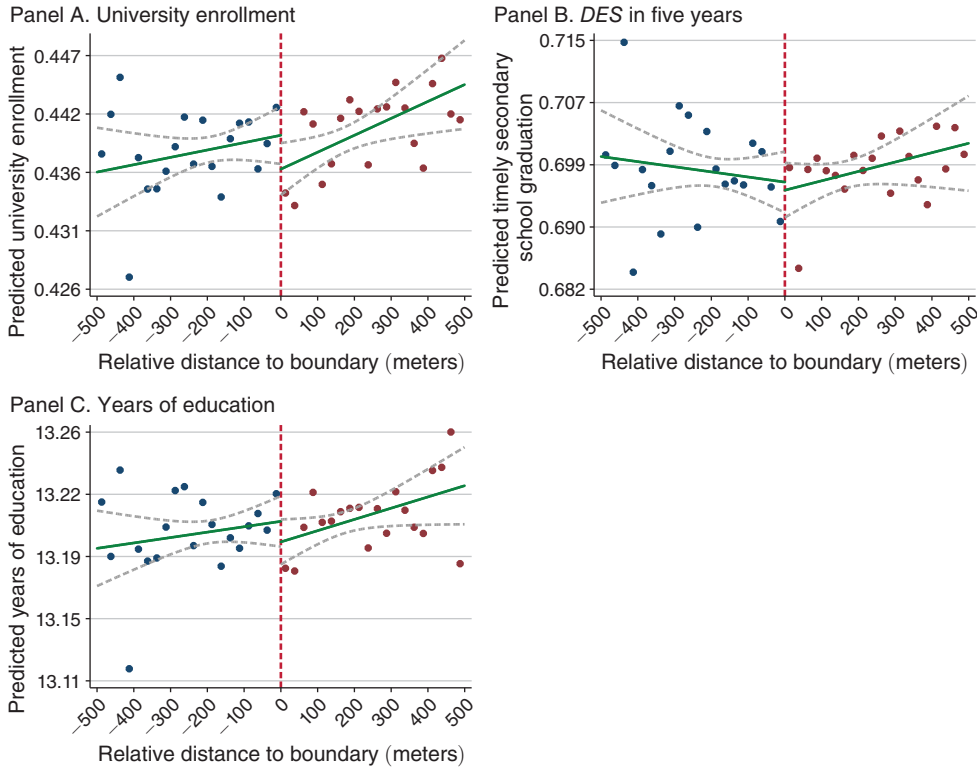


FIGURE 2. BALANCING TEST: PREDICTED EDUCATIONAL ATTAINMENT AT BOUNDARIES

Notes: This figure presents the relationship between predicted educational attainment and distance to the nearest French primary school boundary. Predicted educational attainment is given by the fitted values of a regression of the outcome of interest on individual covariates—gender, place of birth indicators, language at home indicators, use of day care, “in difficulty” status at baseline, handicapped status—and cohort fixed effects. For each boundary, students assigned the default school with the highest predicted gains $\hat{\Omega}_s^P$ are at positive distances. Variables on the vertical axis are residualized on cohort, boundary, and FSA fixed effects. 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. Gray dashed lines are 95 percent confidence intervals based on standard errors clustered at the boundary level.

predicted educational attainment (the fitted values from regressions of educational outcomes on individual covariates X_{icnb}), I find no discontinuity in predicted outcomes. Figure 2 plots predicted outcomes by distance to the nearest boundary, where students assigned to the school of greater quality within a boundary-specific pair are depicted on the right of the threshold (positive distances). For visual clarity, I restrict the sample to permanent residents living within 500 meters of their nearest boundary. For each of the three outcomes, no discontinuity in predicted educational attainment can be discerned. The distribution of each covariate taken on its own also appears to be smooth at the threshold (online Appendix Figures A12; A13; and A14, panels (a) to (j)). Similarly, there is no selective attrition around boundaries (panels (k) and (l)). The associated regression estimates are shown in online Appendix Table A5.

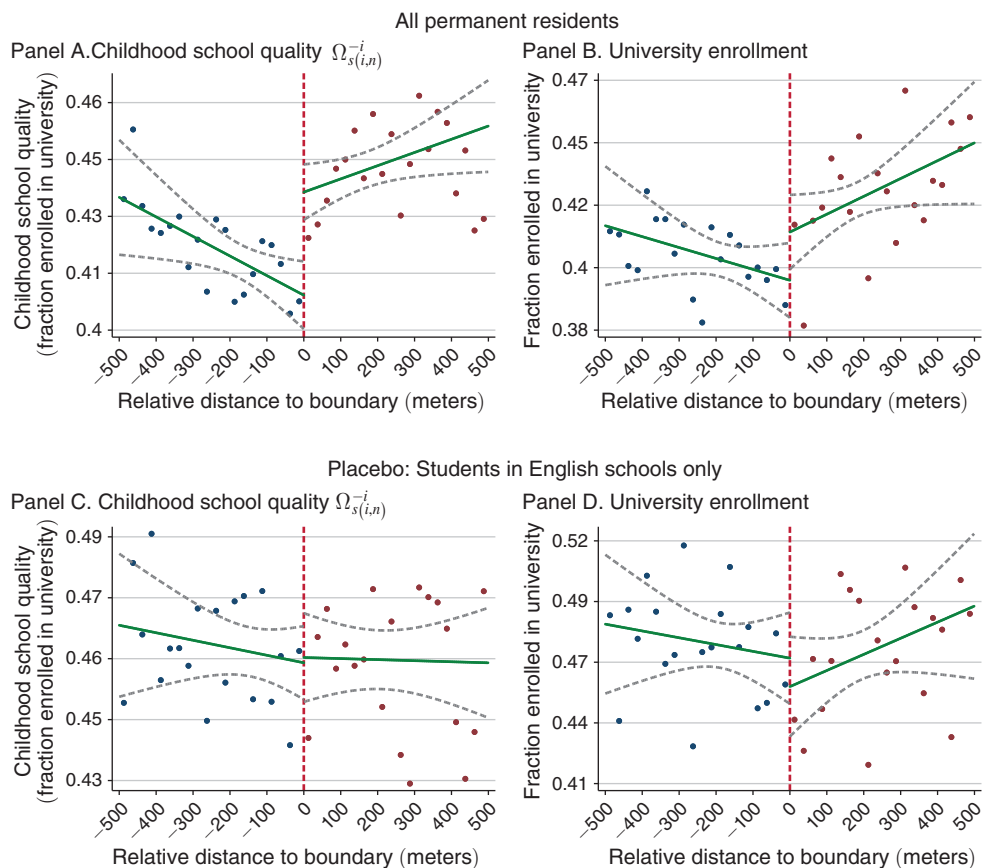


FIGURE 3. REGRESSION-DISCONTINUITY—FIRST-STAGE AND REDUCED-FORM RELATIONSHIPS

Notes: This figure presents the relationship between average leave-self-out childhood school quality $\Omega_{s(i,n)}^{-i}$ (panels A and C) and distance to the nearest French primary school boundary, and between university enrollment and distance to the boundary (panels B and D). For each French primary school boundary, the neighborhood school with greater school quality—in terms of university enrollment—is assigned to the right. 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. Gray dashed lines are 95 percent confidence intervals based on standard errors clustered at the boundary level.

Note that if non-English families were sorting around boundaries on the basis of their willingness to pay for the quality of French schools, any resulting house price response should lead English families to sort in the other direction, as they all participate in the same housing market. Importantly, there is no discontinuity in the fraction of English families around boundaries (panel (f) of Figures A12–A14), consistent with the absence of sorting. In addition, analyses presented in online Appendix D indicate that tract-level house prices are continuously distributed around boundaries.

Results.—Figure 3 shows first-stage and reduced-form relationships between distance to boundaries and university enrollment. Panels A and B include all permanent

TABLE 2—SCHOOL EFFECTS – REGRESSION–DISCONTINUITY ESTIMATES

	First-stage(s)			Reduced- form	RD-IV (π)		
	Dependent variable						
	Quality of assigned school at baseline ($\Omega_{s(i)}^P$)	Quality of school attended at baseline ($\Omega_{s(i)}^P$)	Childhood average school quality ($\Omega_{s(i,n)}^{-i}$)			Parameters	
				Outcome	Outcome	Bandwidth (meters)	Observations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Measure of educational attainment</i>							
All permanent residents							
University enrollment	0.0636 (0.0034)	0.0236 (0.0028)	0.0253 (0.0071)	0.0204 (0.0124)	0.8010 (0.3796)	350	30,067
<i>DES</i> in 5 years	0.0704 (0.0038)	0.0303 (0.0028)	0.0279 (0.0071)	0.0267 (0.0121)	0.9583 (0.3282)	378	31,295
Years of education	0.2915 (0.0152)	0.1106 (0.0125)	0.1091 (0.0347)	0.0940 (0.0589)	0.8551 (0.3892)	332	29,222
Placebo: Students in English schools							
University enrollment	0.0625 (0.0047)	−0.0032 (0.0041)	−0.0014 (0.0096)	−0.0151 (0.0205)	—	358	8,057
<i>DES</i> in 5 years	0.0711 (0.0060)	0.0060 (0.0037)	0.0035 (0.0090)	−0.0008 (0.0207)	—	322	7,527
Years of education	0.2767 (0.0204)	0.0086 (0.0155)	0.0339 (0.0426)	−0.0840 (0.0972)	—	322	7,523
Cohort fixed effects	X	X	X	X	X		
Individual characteristics	X	X	X	X	X		
Neighborhood fixed effects	X	X	X	X	X		
Boundary fixed effects	X	X	X	X	X		

Notes: This table presents regression discontinuity estimates of the causal effect of school quality on educational attainment. Measures of educational attainment are university enrollment, completing secondary school in five years (*DES* in five years) and years of education. Columns 1 through 4 report reduced-form effects of residing on the higher school quality side of a French primary school boundary. The outcome is the quality (fixed effect estimate $\hat{\Omega}_{s(i)}^P$) of the assigned French primary school in column 1, and the quality of the primary school the student actually attends in column 2. In column 3, the dependent variable is the average quality of schools attended during childhood $\Omega_{s(i,n)}^{-i}$ (between the ages of 6 and 15), and in column 4 it is educational attainment y . Column 5 reports 2SLS estimates of the effect of childhood average school quality on educational attainment. Column 6 reports the optimal bandwidth used in the estimation, and column 7 reports the effective number of observations. In all specifications, the control function for distance to boundary is linear and allows for different slopes on either side of the threshold. All models include individual characteristics—gender, place of birth indicators, language at home indicators, use of day care, “in difficulty” status at baseline, handicapped status—as well as cohort, neighborhood (FSA) and boundary fixed effects. 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, shown in parentheses, are clustered at the French primary school boundary level.

residents, and panels C and D are placebo tests restricting the sample to students in English schools. The first graph confirms that the instrument has a strong first-stage. Being assigned a better school at baseline does significantly shift the average quality of schools a student attends during childhood ($\Omega_{s(i,n)}^{-i}$), with school quality increasing with distance on both sides of the cutoff. Educational attainment also jumps right at the threshold: students on the better side of a boundary at age 6 are about 2 percentage points more likely to eventually enroll in university later in life. Importantly, there is no break in school quality or university enrollment for students in English schools. The sharp changes observed at the threshold for the full sample are therefore due to

schools themselves rather than to some other productive neighborhood characteristic that varies discontinuously and coincides with these boundaries.

Regression results analog to Figure 3 are presented in Table 2. Columns 1 through 4 are first-stage and reduced-form regressions and are estimated by OLS. The average quality gap between French default schools on opposite sides of a shared boundary is 6.36 percentage points (SE 0.34) in terms of university enrollment (column 1). The gap is similar for the subsample of English-school students, indicating that they reside around *French* boundaries that are no different than the boundaries faced by the full sample. For all three main outcomes, differences in default options do translate into significant differences in the quality of schools attended at baseline (e.g., gap of 2.36 (SE 0.28) in column 2 for university enrollment). Consistent with the visual evidence, this initial shift in default school quality strongly affects average childhood school quality $\Omega_{s(i,n)}^{-i}$ (column 3). The results in column 4 indicate sizable reduced-form relationships between each measure of educational attainment and the assignment variable. For example, students living on the better side of boundary are 2.67 percentage points (SE 1.21) more likely to obtain a secondary school diploma in five years than students on the opposite side. Crucially, for columns 2 through 4, coefficients for placebo tests reported in the bottom panel are close to zero and statistically indistinguishable from zero.

The last column reports two-stage least square estimates of causal school effects on educational attainment. The RD-IV coefficient of π is below one for university enrollment (0.8010, SE 0.3796) and years of education (0.8551, SE 0.3892), which implies the presence of some degree of sorting into schools that is not accounted for by place of residence. In contrast, the coefficient for finishing secondary school on time is very close to one. Speculatively, for a given amount of sorting, schools likely have a more direct influence on immediate outcomes such as graduating on time than on higher education investments made later in life.

III. Total Exposure Effects

Estimation Framework.—In this section, I estimate the impact of moving to a given neighborhood by investigating whether the educational outcomes of movers converge toward those of the permanent residents of the FSA to which they move in proportion with time spent in that neighborhood. The econometric framework models movers' outcomes as a function of the outcomes of permanent residents of the neighborhoods in which they have resided, weighted by time spent in these locations. In particular, let $\Delta \bar{y}_{od} = \bar{y}_d^{PR} - \bar{y}_o^{PR}$ denote the difference in outcomes between PRs of areas d and o . The main estimating equation is

$$(4) \quad y_{icmod} = \beta(m_i \times \Delta \bar{y}_{od}) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod},$$

where y_{icmod} is some educational outcome of student i , from cohort c , who resided in neighborhood o (origin) at baseline, and moved to neighborhood d (destination) at age m_i . The coefficient of interest β is the annual rate at which outcomes of movers converge to that of the permanent residents in their destination. Positive

exposure effects imply that the cumulative impact of moving to an area where PRs' educational attainment is $\Delta \bar{y}_{od}$ -unit higher should grow (shrink) with the amount of time spent the destination (with age-at-move). Intuitively, if d has greater outcomes than o , then a student who moved at age 9 is expected to have better outcomes than her peer who made the same move at age 12 since she will have been exposed to the destination area for three more years.³¹ The origin is the FSA in which students resided at baseline, while the destination is the one in which they lived during the academic year they were aged 15 on September 30. Sorting into neighborhoods is accounted for by origin-by-destination fixed effects (α_{od}) and unobserved differences between students who move at different ages, notably differential disruption costs, are absorbed by age-at-move fixed effects (α_m). Cohorts fixed effects (α_c) are also included to account for the different number of years for which students are tracked in the data. Standard errors are clustered at the destination neighborhood level to allow for arbitrary correlation among families moving to the same place.³²

To maximize power, in most specifications the sample includes all movers irrespective of the number of times they moved across FSAs, as long as both origin and destination are within Montreal and are not the same. For multiple-times movers, the average quality of neighborhoods exposed to prior to moving to the final destination is therefore measured with error.³³ The model is therefore also estimated on the subsample of one-time movers. In all cases, the sample is always restricted to movers whose origin and destination both have at least 10 permanent residents.

Results.—I begin by providing visual evidence of the convergence of movers' outcomes toward those of the permanent residents of their destination by estimating a semi-parametric version of equation (4). More specifically, in Figure 4, I interact $\Delta \bar{y}_{od}$ with a set of indicators for each possible value of age-at-move m_i (age 7 to 15). As expected, the coefficients on $\Delta \bar{y}_{od}$ shrink (increase) with age-at-move (time spent in destination neighborhood). Importantly, they decrease approximately linearly with age-at-move, which suggests the linear specification in equation (4) is reasonable. The convergence rate β is the *slope* of the line that would best fit the coefficients.

Table 3 reports the results for the main outcomes considered—university enrollment, finishing secondary school on time, and years of education. All models are estimated by ordinary least squares and standard errors are clustered at the destination FSA level. For the two binary outcomes, moving one year earlier to a neighborhood where permanent residents exhibit 10 percentage point higher outcomes, relative to the origin, increases movers' educational attainment by about 0.42 percentage points

³¹ If individual inputs adjust in response to changes in other inputs, then the effect of moving a student across neighborhoods should be interpreted as a policy effect that encompasses parental responses (Todd and Wolpin 2003). For instance, prior research suggests that parental effort and school quality are treated as substitutes (Pop-Eleches and Urquiola 2013, Houtenville and Conway 2008).

³² There is no systematic correlation between m_i and $\Delta \bar{y}_{od}$ in the data. Children who move at early ages are no more likely to move to higher outcome areas (relative to their origin) than children who move at later ages (Figure A15). A Kolmogorov-Smirnov test cannot reject the null of equality of distributions of $\Delta \bar{y}_{od}$ between early (age 7–11) and late (age 12–15) movers (p -value = 0.22).

³³ To later keep the decomposition tractable, I focus on specifications that exploit variation from only two locations (the origin and the destination). Online Appendix Table A7 reports results for models in which I substitute an exposure-weighted average of all locations exposed prior to moving to the final destination for \bar{y}_o^{PR} . About 2/3 of movers move across FSAs only once, and only 6 percent move more than three times.

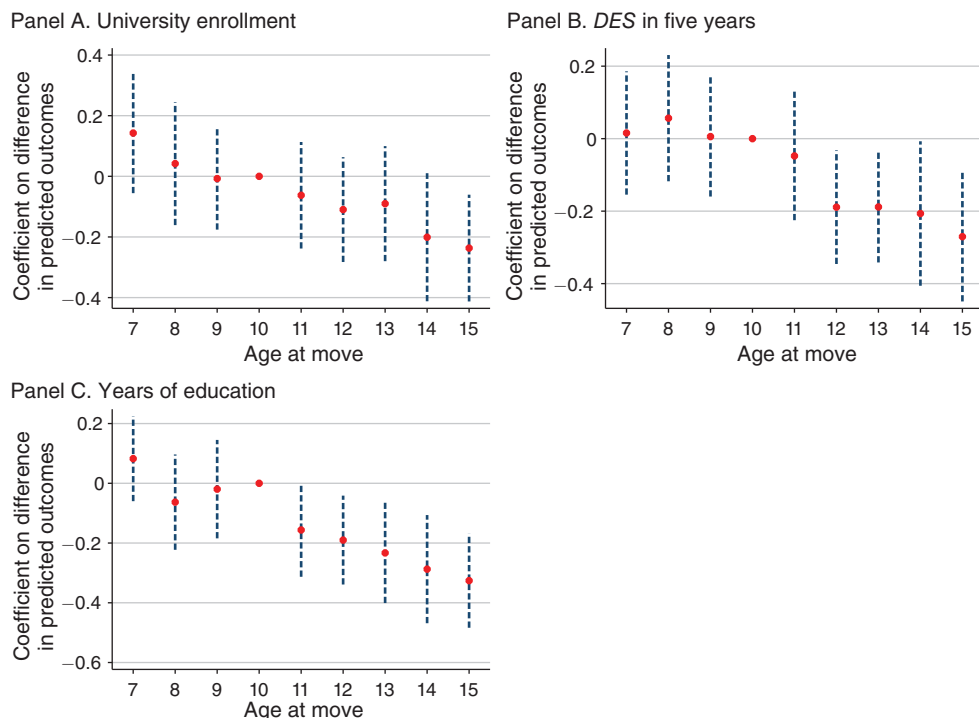


FIGURE 4. SEMI-PARAMETRIC TOTAL EXPOSURE EFFECTS

Notes: This figure presents regression estimates of the effect of moving to a given area on educational attainment (total exposure effects). In particular, it plots estimates of coefficients β_m against age-at-move m , from the estimating equation $y_{icmod} = \sum_{m=7}^{15} \beta_m (\Delta \bar{y}_{od} \times \mathbf{1}\{m_i = m\}) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \epsilon_{icmod}$, where i indexes a student, o the origin FSA, d the destination FSA, and c cohorts. The main regressors are interactions between $\Delta \bar{y}_{od} = \bar{y}_d - \bar{y}_o$ —the difference in outcomes between permanent residents of areas d and o —and indicators for age-at-move. The sample includes all movers who remained within Montreal. All coefficients β_m are relative to moving at age 10. In panel A, the outcome y is university enrollment, in panel B it is an indicator for finishing secondary school within five years, and in panel C it is the number of years of education. 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 percent confidence intervals for the point estimates.

(column 1). Extrapolating over 10 years, the cumulative effect would therefore be 4.2 percentage points, or 42 percent of the difference between permanent residents of the destination and origin locations. These point estimates are all statistically significant at the 1 percent level. A slightly larger coefficient is obtained for years of education, implying a convergence rate of about 4.9 percent. Controlling for poor schooling outcomes prior to moving (dummies for the number of times in difficulty, columns 2 and 4) or restricting the sample to one-time movers (columns 3 and 4) barely affects the magnitudes of the coefficients. In online Appendix Table A8, I further include 6-digit postal code fixed effects, thereby restricting the comparison between children who at the age of 15 lived either on the same block or in the same apartment building, as an attempt to absorb as much variation in movers' parental income as possible. The fact that many postal codes contain only one observation shrinks the sample size substantially, but the estimated convergence rates remain

TABLE 3—TOTAL EXPOSURE EFFECTS—MOVERS ESTIMATES

Sample:	All movers		One-time movers	
	(1)	(2)	(3)	(4)
<i>Measure of educational attainment</i>				
University enrollment	−0.0424 (0.0090)	−0.0412 (0.0092)	−0.0416 (0.0116)	−0.0408 (0.0115)
Secondary school diploma in five years	−0.0421 (0.0088)	−0.0402 (0.0088)	−0.0506 (0.0117)	−0.0502 (0.0117)
Years of education	−0.0488 (0.0088)	−0.0471 (0.0094)	−0.0444 (0.0103)	−0.0435 (0.0102)
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
Observations	24,316	24,316	15,533	15,533

Notes: This table reports regression estimates of total exposure effects on educational attainment. Estimates of the convergence rate β are estimated from a regression of movers outcomes on $\Delta \bar{y}_{od} = \bar{y}_d - \bar{y}_o$, the difference in outcomes between permanent residents of the destination and origin neighborhoods, interacted with age-at-move (equation (4)). The coefficients can be interpreted as the impact of moving one year later to a neighborhood where permanent residents have one-unit greater educational outcomes. In columns 1 and 2, all movers are included, whereas in columns 3 and 4 the sample is restricted to one-time movers. Students who moved from or to neighborhoods (FSAs) with less than 10 permanent residents are always excluded. All regressions include age-at-move, origin-by-destination, and cohort fixed effects. Individual characteristics are gender, immigrant status, allophone status, born in Canada but outside Quebec, English spoken at home, day care use at baseline, “in difficulty” status at baseline, handicapped status. In columns 2 and 4, the model includes a set of dummies for each possible value of number of times a student is identified as being “in difficulty” prior to moving. Standard errors, shown in parentheses, are clustered at the destination neighborhood level. The movers sample contains a total of 25,993 observations, of which 1,677 are singletons and therefore dropped in the estimation.

qualitatively similar to the benchmark estimates. Results for alternative measures of educational attainment mirror the main results (online Appendix Table A6).³⁴

My estimates of total exposure effect are surprisingly close to those reported by Chetty and Hendren (2018a), who find a convergence rate of 4 percent in earnings for millions of movers across commuting zones in the United States. While one may expect a larger influence of neighborhoods at finer levels of geography, my estimates are more likely to be attenuated due to sampling error in the calculation of the average outcomes of permanent residents. Nonetheless, it is remarkable that our findings so closely align given the differences in the locations, populations and outcomes we study. While movers *across* cities tend to have a slight income advantage relative to stayers, movers *within* Montreal appear to be negatively selected.

The main specification not only assumes that exposure effects are linear with age-at-move, but also that they are linear and symmetric in $\Delta \bar{y}_{od}$. I explore the validity of this assumption by estimating convergence rates separately for students who move to higher and lower educational attainment areas. While the point estimates

³⁴ Allowing permanent residents’ outcomes to be cohort-specific (\bar{y}_{hc}^{PR}) increases sampling error and therefore yields convergence rates of slightly smaller magnitudes (3.4–4.3 percent). Similar patterns emerge if I use mutually exclusive cohorts to calculate \bar{y}_n^{PR} and to implement the movers design. These results are available upon request.

differ for positive and negative moves, I cannot reject that the two coefficients are statistically equal at conventional levels (online Appendix Table A9, columns 5 and 6).

Another possible concern is that my preferred definition of neighborhoods might be too large to accurately capture the social context children are exposed to outside of school. Online Appendix E considers census tracts as an alternative unit of analysis to examine whether the results are affected by the choice of geography. Convergence rates are smaller at the census tract level than at the FSA level, possibly because of larger sampling error and of greater sorting of permanent residents at smaller levels of geography.³⁵ Census tracts may also less precisely reflect all features of the community in which children live and socialize, which is arguably larger than a single census tract. In fact, I find that any benefits of moving to a higher educational attainment census tract are driven by between-FSA moves—no convergence is found for moves that occur within FSAs but across census tracts.

Estimates of the total effect of one year of exposure to a given area on educational attainment are valid under the assumption that the degree of selection into FSAs does not vary systematically with age. In Section VI, I conduct a host of robustness checks, including family fixed-effects models, to corroborate the validity of this assumption.

IV. Conceptual Decomposition Framework

This section lays out a human capital production function that linearly incorporates cumulative school and neighborhood non-school inputs, expanding the framework of Chetty and Hendren (2018b) to allow for multiple contextual dimensions. It first describes the outcomes of permanent residents and movers parsimoniously in terms of school and neighborhood quality. These expressions map to the reduced-form parameters previously estimated. Finally, I discuss how a decomposition of total exposure effects is achieved.

Education Production Function.—Consider a framework in which educational investment in children takes place over compulsory schooling years (up to year A) and a long-term outcome is realized and measured after the investment years. The education production function is cumulative and separately additive in yearly family, school, and neighborhood (non-school) inputs:

$$(5) \quad y_i = \sum_n a_{in} [\lambda \mu_n + \omega \tilde{\psi}_{s(i,n)}] + A \tilde{\theta}_i,$$

where y_i is a measure of educational attainment for child i , a_{in} is the number of years the child resided in location n (with $\sum_n a_{in} = A$), $\tilde{\psi}_{s(i,n)}$ denotes the annual average quality of schools attended by child i while residing in location n , and $\tilde{\theta}_i$ are annual

³⁵ The converge rate varies between 2 percent and 2.5 percent at the census tract level. Chetty et al. (2018) also obtain a rate of 2.5 percent at the census tract level in the United States.

average family inputs.³⁶ Neighborhood and school quality are denoted by variables μ_n and $\tilde{\psi}_{s(i,n)}$, and parameters λ and ω respectively represent the causal effect of one year of exposure to non-school neighborhood amenities and the causal effect of attending a given school for one year.³⁷ Average school quality $\tilde{\psi}_{s(i,n)}$ remains indexed by i because students living in the same area can attend different schools.

This education production function has several important restrictions. First, neighborhood and school effects are both linear in years of exposure.³⁸ Figure 4 suggests that movers' outcomes indeed converge linearly to those of permanent residents of the destination. Second, the model imposes additive separability of schools and neighborhoods, thereby ruling out complementarities between the two dimensions. Evidence that there is no systematic interaction between school and neighborhood quality is provided in Section ID. Additive separability of schools and neighborhoods is also relatively standard in the literature (Gibbons, Silva, and Weinhard 2013; Card and Rothstein 2007), and is consistent with results from the Harlem Children's Zone (Fryer and Katz 2013).

Importantly, school and neighborhood effects are assumed to be constant across students, an assumption that is also common to most work on school (Deming 2014), teacher (Chetty, Friedman, and Rockoff 2014) and college (Hoxby 2015) value-added. This assumption has two key implications. First, it insures that local RD estimates of the effect of school quality can be generalized to the overall student population. I examine this assumption empirically in Section VIA. Second, it guarantees contextual effects are the same for movers and permanent residents.³⁹

In contrast, Chetty and Hendren (2018a) allow contextual effects to vary by income level. Empirically, one concern is that since permanent residents and movers likely exhibit systematic income differences, these two groups might be subject to different neighborhood and school effects, resulting in downward biased estimated total exposure effects. I do not observe parental income, but provide some evidence suggesting movers and permanent residents benefit from school and neighborhood quality similarly. Online Appendix Figure A16 (left panel) plots neighborhood-level average outcomes for permanent residents against

³⁶ School s 's quality is given by ψ_s , and $\tilde{\psi}_{s(i,n)}$ is the student-specific average of ψ_s taken across years during which student i resided in location n .

³⁷ To keep the model tractable, I do not explicitly include disruption costs associated with moving or switching school in the production function. The movers design accounts for any age-variant disruption costs with the inclusion of age-at-move fixed effects. Also, the framework abstracts from school and neighborhood effects that are not proportional with length of exposure. For instance, any benefits associated with residing close to a university at the age of 20, which are unrelated to how old one was when one arrived at that location, are not taken into account. This is because any such effects are absorbed by origin-by-destination fixed effects when estimating total exposure effects, and the inclusion of these fixed effects is necessary for identification. However, such exposure-independent effects may not be that large in a city like Montreal, where numerous universities and colleges are scattered across the city. In addition, Chetty and Hendren (2018a) find that place effects converge toward zero as age-at-move approaches 23, and Chetty, Hendren, and Katz (2016) find no benefits of moving to a low-poverty area in adulthood in the MTO experiment. Further details on the production function are provided in online Appendix C.

³⁸ Angrist et al. (2017), Dobbie and Fryer (2013), Abdulkadiroğlu et al. (2011), and Autor et al. (2016) also assume that school effects are proportional with number of years. Chetty and Hendren (2018a) make a similar assumption for place effects.

³⁹ Alternatively, if neighborhoods and schools have heterogeneous effects, μ_n and ψ_s can be conceptualized as average effects. The decomposition approach then explicitly requires the assumptions that (i) the local weighted average of school effects from the RD design can be generalized to the broader student population, and (ii) that average school and neighborhood effects are the same for movers and permanent residents.

exposure-weighted average outcomes for movers. The raw correlation between the two averages is very high, around 0.92, and consistent with a signal correlation close to 1 given the reliability ratios reported in Table 1. The attainment gap between permanent residents and movers does not appear to vary systematically with neighborhood outcomes. In other words, the difference in mean outcomes between any two neighborhoods is roughly the same for both groups throughout the distribution. In addition, gains from moving are symmetric: neighborhoods that produce high outcomes for movers who originated in them produce similar outcomes for movers who chose them as their destination (online Appendix Figure A16, right panel). In Section VIC, I additionally verify that directly controlling for the schools attended by movers in equation (4) produces results almost identical to those obtained by projecting permanent residents' estimated school effects onto movers.

School Effects.—In practice, the terms $\tilde{\psi}_{s(i,n)}$ and μ_n are unobserved. To obtain measures of school quality, I focus on the subsample of permanent residents (PR)—children who always resided in the same place k —for whom $a_{in} = A$ for neighborhood k , and $a_{in} = 0$ for all other locations $n \neq k$. For these students, $y_i^{PR} = A[\lambda\mu_n + \omega\tilde{\psi}_{s(i,n)} + \tilde{\theta}_i]$. I partition their educational attainment into measurable school-related and neighborhood-related terms, and an idiosyncratic residual ν_i unrelated to either schools or neighborhoods:

$$(6) \quad y_i^{PR} = \Omega_{s(i,n)} + \Lambda_n + \nu_i,$$

where $\Omega_{s(i,n)}$ reflects both cumulative causal school effects over student i 's childhood $A\omega\tilde{\psi}_{s(i,n)}$ as well as average sorting into schools, and Λ_n is defined accordingly for neighborhood non-school amenities. Put differently, estimated school quality $\Omega_{s(i,n)}$ might be a biased measure of true school effects because it also incorporates within-neighborhood school-level differences in parental inputs. In practice, I estimate $\Omega_{s(i,n)}$ and Λ_n using a two-way fixed effects model (equation (1)).

My first reduced-form parameter of interest is the average causal effect of an increase in school quality $\Omega_{s(i,n)}$ on educational outcomes. Let π denote the coefficient from the unfeasible regression of causal school effects $A\omega\tilde{\psi}_{s(i,n)}$ on estimated school quality $\Omega_{s(i,n)}$. Hence, $1 - \pi$ is the amount of forecast-bias in estimates of $\Omega_{s(i,n)}$.⁴⁰ Estimation of π can be achieved using a valid instrumental variable that exogenously shifts school quality independently of neighborhood quality (first-stage) and that is uncorrelated with parental inputs (exclusion restriction). Whereas an OLS regression of y_i^{PR} on $\Omega_{s(i,n)}$ and a set of neighborhood fixed effects yields a coefficient on $\Omega_{s(i,n)}$ of one, by construction, the RD-IV design used in Section II isolates causal variation in estimated school quality and thereby pins down π . Letting \bar{x}_k^{PR} denote the average of some variable x for PRs of neighborhood k , $\pi\bar{\Omega}_k^{PR}$ then provides a forecast-unbiased measure of the average cumulative causal school effects for PRs of neighborhood k , $A\omega\bar{\psi}_k^{PR}$.

⁴⁰ More details about the interpretation of π are provided in online Appendix C.

Total Exposure Effects.—For one-time movers, let $o(i)$ denote the origin neighborhood of mover i , $d(i)$ denote the destination, and m_i the age at which student i moved. For these students, $a_{in} = m_i - 1$ for $n = o$, $a_{in} = A - (m_i - 1)$ for $n = d$, and $a_{in} = 0$ for all $n \neq o, d$. Their educational attainment is given by

$$(7) \quad y_i = \alpha_{od} - (m_i - 1) \underbrace{\left[\lambda(\mu_d - \mu_o) + \omega(\tilde{\psi}_{s(i,d)} - \tilde{\psi}_{s(i,o)}) \right]}_{\text{Total exposure effects } (e_{i,od})} + \epsilon_i,$$

where $\alpha_{od} = A\lambda\mu_d + A\omega\bar{\psi}_{s(d)} + A\bar{\theta}_{od}$ captures what is common to students who made the move from o to d , and $\epsilon_i = A(\bar{\theta}_i - \bar{\theta}_{od}) + A(\omega\tilde{\psi}_{s(i,d)} - \omega\bar{\psi}_{s(d)})$ is student-level idiosyncratic variation. Equation (7) highlights that the long-term outcomes of movers depend on the quality of schools and neighborhoods in both places as well as on the length of *exposure* to each place, which varies with age-at-move. Total exposure effects $e_{i,od}$ are the gains of living in and attending schools of area d for one year relative to area o . Under the constant effects assumption, school and neighborhood effects are the same for movers and PRs and one can rewrite $e_{i,od}$ as a function of the difference in PR mean outcomes $\Delta\bar{y}_{od}$:

$$(8) \quad e_{i,od} = \left(\frac{1}{A}\right)\Delta\bar{y}_{od} + (c_{i(od)} - 1)\omega[\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR}] - [\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}],$$

where $c_{i,od} = (\tilde{\psi}_{s(i,d)} - \tilde{\psi}_{s(i,o)})/(\bar{\psi}_d^{PR} - \bar{\psi}_o^{PR})$ is a school compliance factor indicating the propensity of movers to attend schools of comparable quality to those attended by permanent residents in their origin and destination.

Let β denote the coefficient from the unfeasible regression of total exposure effects $e_{i,od}$ on estimated changes in neighborhood quality $\Delta\bar{y}_{od}$ and $r_i = e_{i,od} - \beta\Delta\bar{y}_{od}$ the associated residual. Movers' outcomes can then be written

$$(9) \quad y_i = \alpha_{od} - \beta(m_i - 1)\Delta\bar{y}_{od} + u_i,$$

where $u_i = (m_i - 1)r_i + \epsilon_i$. The reduced-form parameter β is the causal effect of moving to an area where PRs have $\Delta\bar{y}_{od}$ -unit higher outcomes and can be interpreted as a convergence rate. If permanent residents were not to sort across neighborhoods ($\bar{\theta}_o^{PR} = \bar{\theta}_d^{PR} \forall o, d$) and movers always attended the same schools as permanent residents ($c_{i,od} = 1 \forall i$), then movers outcomes would converge toward those of PRs at an annual rate of $\beta = 1/A$. Otherwise, β is increasing in $c_{i,od}$ and decreasing in the degree of sorting of permanent residents.⁴¹ If neither schools nor neighborhoods matter (i.e., $\lambda = \omega = 0$) and $\Delta\bar{y}_{od}$ solely reflects differences in parental inputs, then $\beta = 0$.

⁴¹ This is under the assumption that families with high unobservable characteristics select into schools and neighborhoods that produce higher outcomes: $\text{cov}(\lambda\mu_n + \omega\bar{\psi}_n^{PR}, \bar{\theta}_n^{PR}) > 0$. In contrast, if low $\bar{\theta}_i$ families are more likely to sort into such schools and neighborhoods, the convergence rate increases with sorting of permanent residents.

Decomposition.—Moving to a new neighborhood produces changes in school and non-school neighborhood inputs. Here, I decompose the total effect of moving to a given neighborhood into a school component reflecting differences in causal schools effects across locations and a non-school component attributable to other neighborhood amenities. Among permanent residents, spatial differences in estimated school quality cause educational attainment to be higher by $\pi\Delta\Omega_{od} = \pi(\bar{\Omega}_d^{PR} - \bar{\Omega}_o^{PR})$ in location d relative to o , on average. Taking out these differences, the remaining gap, $\Delta\bar{y}_{od}^{-s} \equiv \Delta\bar{y}_{od} - \pi\Delta\Omega_{od}$, reflects non-school factors. Assuming that the causal effect of school quality π is the same for movers and permanent residents,⁴² one can decompose total exposure effects accordingly:

$$(10) \quad e_{i,od} = \left(\frac{1}{A}\right)\Delta\bar{y}_{od}^{-s} + \left(\frac{c_{i,od}}{A}\right)\pi\Delta\Omega_{od} - [\bar{\theta}_d^{PR} - \bar{\theta}_o^{PR}] + v_i,$$

where v_i captures idiosyncratic deviations in school quality from their unbiased forecast $\pi\Delta\Omega_{od}$. From an accounting perspective, the total convergence rate β can similarly be separated into school and non-school components:

$$(11) \quad \beta = \frac{\text{cov}(e_{i,od}, \Delta\bar{y}_{od})}{\text{var}(\Delta\bar{y}_{od})} = \underbrace{\frac{\text{cov}(e_{i,od}, \Delta\bar{y}_{od}^{-s})}{\text{var}(\Delta\bar{y}_{od}^{-s})} \frac{\text{var}(\Delta\bar{y}_{od}^{-s})}{\text{var}(\Delta\bar{y}_{od})}}_{\beta^{non-school}} + \underbrace{\frac{\text{cov}(e_{i,od}, \pi\Delta\Omega_{od})}{\text{var}(\pi\Delta\Omega_{od})} \frac{\text{var}(\pi\Delta\Omega_{od})}{\text{var}(\Delta\bar{y}_{od})}}_{\beta^{school}}.$$

For instance, the non-school component $\beta^{non-school}$ is equal to the coefficient from the unfeasible regression of $e_{i,od}$ on $\Delta\bar{y}_{od}^{-s}$ —the effect of moving to a place that has better outcomes for non-school reasons—scaled by $\text{var}(\Delta\bar{y}_{od}^{-s})/\text{var}(\Delta\bar{y}_{od})$ to insure it has the same denominator as the full convergence rate. I define the share of total effects accounted for by causal school (non-school) effects to be the ratio of the school (non-school) component over the total convergence rate

$$(12) \quad s^{school} \equiv \frac{\beta^{school}}{\beta} = \frac{\text{cov}(e_{i,od}, \pi\Delta\Omega_{od})}{\text{cov}(e_{i,od}, \Delta\bar{y}_{od})},$$

$$s^{non-school} \equiv \frac{\beta^{non-school}}{\beta} = \frac{\text{cov}(e_{i,od}, \Delta\bar{y}_{od}^{-s})}{\text{cov}(e_{i,od}, \Delta\bar{y}_{od})}.$$

The *school share's* magnitude decreases with the amount of sorting of permanent residents into schools. It increases with the compliance factor $c_{i,od}$ and therefore depends on the school choice regime analyzed. Also, s^{school} increases linearly with π , and is effectively equal to zero if $\pi = 0$. Under full school compliance ($c_{i,od} = 1 \forall i$) and no sorting of permanent residents, the school share is

⁴²In section VIA, I use a model-based approach to estimate π for movers and find values in line with those for permanent residents.

the fraction of the total variance $\text{var}(\Delta \bar{y}_{od})$ that is attributable to schools, that is $\text{cov}(\Delta \Omega_{od}, \Delta \bar{y}_{od}) / \text{var}(\Delta \bar{y}_{od}) = (\text{var}(\Delta \Omega_{od}) + \text{cov}(\Delta \Omega_{od}, \Delta \bar{y}_{od}^{-s})) / \text{var}(\Delta \bar{y}_{od})$.

V. Decomposition: Schools or Neighborhoods?

Estimation Framework.—The total convergence rate β reflects the combined effect of changes in school and neighborhood (non-school) quality. To quantify the role of schools as a driver of this total effect, I reestimate the movers design (equation (4)) substituting $(m_i \times \Delta \bar{y}_{od}^{-s})$ and $(m_i \times \pi \Delta \Omega_{od})$ for $(m_i \times \Delta \bar{y}_{od})$:

$$(13) \quad y_{icmod} = \beta_s(m_i \times \pi \Delta \Omega_{od}) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod},$$

$$(14) \quad y_{icmod} = \beta_n(m_i \times \Delta \bar{y}_{od}^{-s}) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod}.$$

The school and non-school components of total exposure effects are calculated as

$$(15) \quad \beta^{school} = \beta_s \frac{\text{var}^r(\pi \Delta \Omega_{od})}{\text{var}^r(\Delta \bar{y}_{od})}; \quad \beta^{non-school} = \beta_n \frac{\text{var}^r(\Delta \bar{y}_{od}^{-s})}{\text{var}^r(\Delta \bar{y}_{od})},$$

where $\text{var}^r(z)$ denotes the variance of the residuals of $(m_i \times z)$.⁴³ The school and non-school shares, S^{school} and $S^{non-school}$, are then obtained by dividing the school and non-school components by the total convergence rate β . Intuitively, $S^{non-school}$ indicates the fraction of total gains that remains after taking out predicted differences in outcomes between origin and destination on the basis of differences in average forecast-unbiased school effects. In online Appendix C.3, I present an alternative but numerically equivalent approach to estimating the parameters β^{school} and $\beta^{non-school}$ from a “horse-race”-type model that simultaneously includes changes in both components of permanent residents’ outcomes, $(m_i \times \pi \Delta \Omega_{od})$ and $(m_i \times \Delta \bar{y}_{od}^{-s})$, in the movers design.

Results.—Again, I start by presenting visual evidence based on semi-parametric estimates. Figure 5 reproduces in light gray the total exposure regression coefficients that were previously shown, and displays the corresponding $\beta^{non-school}$ components in red. For all three outcomes, the slope of the line that connects these points is considerably flatter, indicating a much lower rate of convergence once school effects have been suppressed.

Estimates of β^{school} , $\beta^{non-school}$, and S^{school} are presented in Table 4. Columns 1 through 3 decompose total exposure effects for all movers, while columns 4 through 6 focus on one-time movers. In the first column, π is naïvely set to 1 and school quality is measured by the simple FSA-level average of the sum of primary and secondary school fixed effects ($\bar{\Omega}_n^{PR}$). The non-school component $\beta^{non-school}$ is 0.010 (SE 0.004) for university enrollment, 0.012 (SE 0.004) for finishing secondary school on time, and 0.014 (SE 0.003) for years of education. The corresponding

⁴³For instance, $\text{var}^r(\Delta \bar{y}_{od})$ is the variance of the residuals of the following regression: $(m_i \times \Delta \bar{y}_{od}) = \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod}$.

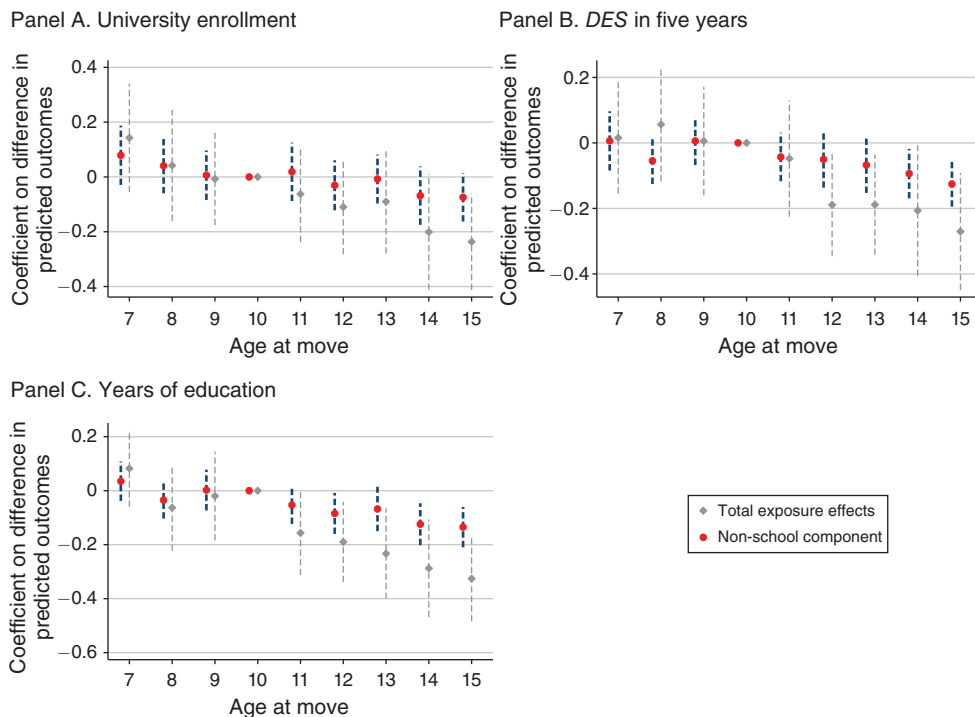


FIGURE 5. SEMI-PARAMETRIC EXPOSURE EFFECTS—NON-SCHOOL COMPONENT

Notes: This figure plots estimates of the non-school components $\beta_m^{non-school}$ against age-at-move m (in red). These estimates reflect the part of the effect of moving to a FSA where permanent residents have greater educational attainment that remains after subtracting spatial differences in school effects. Diamonds (light grey) reproduce the total exposure effects estimates shown in Figure 4. The estimating equation is $y_{icmod} = \sum_{m=7}^{15} \beta_{n,m}(\Delta \bar{y}_{od}^s \times \mathbf{1}\{m_i = m\}) + \gamma X_{icmod} + \alpha_{od} + \alpha_m + \alpha_c + \epsilon_{icmod}$, where i indexes a student, o the origin FSA, d the destination FSA, and c cohorts, and components $\beta_m^{non-school}$ are given by $\beta_{n,m}(\text{var}'(\Delta \bar{y}_{od}^s) / \text{var}'(\Delta \bar{y}_{od}))$. The main regressors are interactions between $\Delta \bar{y}_{od}^s$ —the residual difference in outcomes between permanent residents of areas d and o that remains after subtracting differences in school effects—and indicators for age-at-move. In panel A, the outcome y is university enrollment, in panel B it is an indicator for finishing secondary school within five 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 percent confidence intervals for the point estimates.

school shares S^{school} , which reflect any variation associated with school attendance, are 76 percent, 72 percent, and 72 percent, respectively. In column 2, FSA-level average school quality is measured using permanent residents' leave-self-out childhood school quality $\bar{\Omega}_n^i$, the variable used to estimate π . This minor change in measurement has little effect on the results—the fraction of total exposure effects on educational attainment explained by schools remains in the vicinity of 70–75 percent.

Decompositions based on biased measures of school quality may overstate the importance of schools. My preferred specification, in column (3), uses forecast-unbiased school quality $\pi \bar{\Omega}_n^i$. Here, the school share drops to 60 percent for university enrollment, to 66 percent for finishing secondary school on time, and to 60 percent for years of education. The pattern is similar for one-time movers, with school shares ranging between 61 percent and 71 percent in column (6). The

TABLE 4—DECOMPOSITION OF TOTAL EXPOSURE EFFECTS

Sample:	All movers			One-time movers		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>University enrollment</i>						
Total exposure effects						
β	−0.0424 (0.0090)	−0.0424 (0.0090)	−0.0424 (0.0090)	−0.0416 (0.0116)	−0.0416 (0.0116)	−0.0416 (0.0116)
School and non-school components						
β^{school}	−0.0321 (0.0071)	−0.0320 (0.0070)	−0.0257 (0.0056)	−0.0320 (0.0092)	−0.0318 (0.0090)	−0.0255 (0.0072)
$\beta^{non-school}$	−0.0103 (0.0038)	−0.0104 (0.0038)	−0.0168 (0.0045)	−0.0096 (0.0058)	−0.0097 (0.0059)	−0.0161 (0.0065)
School share (S^{school})	76% (0.0678)	76% (0.0678)	60% (0.0543)	77% (0.1125)	77% (0.1124)	61% (0.0901)
<i>Secondary school diploma in five years</i>						
Total exposure effects						
β	−0.0421 (0.0088)	−0.0421 (0.0088)	−0.0421 (0.0088)	−0.0506 (0.0117)	−0.0506 (0.0117)	−0.0506 (0.0117)
School and non-school components						
β^{school}	−0.0305 (0.0083)	−0.0290 (0.0079)	−0.0278 (0.0076)	−0.0398 (0.0106)	−0.0373 (0.0104)	−0.0358 (0.0100)
$\beta^{non-school}$	−0.0116 (0.0039)	−0.0131 (0.0041)	−0.0143 (0.0041)	−0.0108 (0.0049)	−0.0133 (0.0050)	−0.0148 (0.0050)
School share (S^{school})	72% (0.0918)	69% (0.0905)	66% (0.0867)	79% (0.0889)	74% (0.0890)	71% (0.0853)
<i>Years of education</i>						
Total exposure effects						
β	−0.0488 (0.0088)	−0.0488 (0.0088)	−0.0488 (0.0088)	−0.0444 (0.0103)	−0.0444 (0.0103)	−0.0444 (0.0103)
School and non-school components						
β^{school}	−0.0350 (0.0075)	−0.0340 (0.0072)	−0.0290 (0.0061)	−0.0328 (0.0088)	−0.0316 (0.0085)	−0.0271 (0.0072)
$\beta^{non-school}$	−0.0138 (0.0034)	−0.0148 (0.0035)	−0.0198 (0.0039)	−0.0115 (0.0046)	−0.0127 (0.0046)	−0.0173 (0.0050)
School share (S^{school})	72% (0.0597)	70% (0.0570)	60% (0.0487)	74% (0.0875)	71% (0.0851)	61% (0.0728)
Measure of school quality						
π	$\pi \bar{\Omega}_n$ 1	$\pi \bar{\Omega}_n^{-i}$ 1	$\pi \bar{\Omega}_n^{-i}$ RD estimate	$\pi \bar{\Omega}_n$ 1	$\pi \bar{\Omega}_n^{-i}$ 1	$\pi \bar{\Omega}_n^{-i}$ RD estimate

Notes: This table reports estimates of the share of total exposure effects that is due to spatial differences in school effects. It presents 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. $\beta^{non-school}$ can be interpreted as the part of the impact of moving to a neighborhood where permanent residents have greater educational attainment that remains after subtracting spatial differences in school effects. In columns 1 to 3, all movers are included in the estimating sample, whereas in columns 4 to 6 the sample is restricted to one-time movers. Students who moved from or to neighborhoods (FSAs) with less than 10 permanent residents are always excluded. In columns 1 and 4, 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 FSA-level averages of permanent residents' leave-self-out childhood school quality $\bar{\Omega}_n^{-i}$. In columns 1 and 2, the causal effect of school quality π is set to one. In column 3, π is given by RD-IV estimates reported in Table 2. Standard errors, shown in parentheses, are clustered at the destination FSA level, and obtained by the delta method. Reported total exposure effects β correspond to estimates reported in odd numbered columns in Table 3.

values of β^{school} reported in column (6) suggest that movers' outcomes would converge at rates of 2.6 percent (SE 0.0072), 3.6 percent (SE 0.0100), and 2.7 percent (SE 0.0072) toward mean outcomes of permanent residents for university

enrollment, timely secondary school graduation, and years of education, respectively, on the basis of average differences in causal school effects alone. Using values of π based on different RD bandwidths doesn't affect the results materially, with the vast majority of estimates of the school share S^{school} falling between 50 percent and 70 percent (online Appendix Figure A18).

In online Appendix E, the decomposition exercise is implemented using census tracts instead of FSAs. School shares S^{school} here range from 44 percent for university enrollment to 67 percent for timely graduation from secondary school. In comparison, Chetty et al. (2018) find that 51 percent of the within-county, between census tracts variation in upward mobility is accounted for by high school catchment areas in the United States. While school quality may play a greater role for educational attainment than it does for income mobility, the discrepancy between Chetty et al. (2018) and my decomposition results may also be due to methodological differences. For instance, some students living in the same catchment area may attend schools of different quality. While my methods do take these enrollment patterns into account, Chetty et al. (2018) do not observe actual school enrollment. Also, the Opportunity Atlas focuses on high school catchment areas, whereas I also consider primary school attendance, for which the catchment areas are smaller. Replicating the analysis of Chetty et al. (2018) in my data, I find that about 50 percent to 55 percent of the between census tract variation in \bar{y}_n^{PR} is accounted for by French high school catchment areas, which is very similar to their findings.⁴⁴

Overall, the decomposition analysis indicates that schools matter more than neighborhoods for long-term educational attainment, with most estimates that account for bias in measures of school quality ranging between 50 percent and 70 percent. Nonetheless, schools do not *fully* account for total exposure effects—neighborhoods do have a small independent effect on human capital accumulation. Note that the school contribution to total exposure effects on educational attainment may be larger in contexts where school choice is more restricted, since the school share S^{school} increases with the compliance factor $c_{i,od}$.

VI. Robustness

A. Regression-Discontinuity Estimates

The benchmark specification for estimating school effects imposes several restrictions. First, it assumes that the relationship between distance to the boundary and student outcomes is linear. Online Appendix Table A10 allows for a global quadratic functional form. RD-IV estimates of π with quadratic functions slightly smaller than the baseline for university enrollment (0.71) and years of education (0.80), but greater for finishing secondary school on time (1.08). Using a triangular kernel for the linear control function yields results almost identical to the baseline (online Appendix Table A11).

⁴⁴ These are adjusted R^2 s from regressions of \bar{y}_n^{PR} (at the census tract level) on default French high school fixed effects. If one uses default primary school fixed effects instead, the adjusted R^2 s increase to 76–79 percent.

Sensitivity to bandwidth restrictions is examined in online Appendix Figure A17. Moving along the horizontal axis, I gradually expand the sample by including students living farther away from boundaries. The point estimates fluctuate across sample restrictions, following no monotonic pattern, but gains in precision are achieved with greater bandwidths. For instance, for university enrollment, keeping only students living within 250 meters of a boundary yields a relatively smaller π coefficient of 0.72, while further expanding the bandwidth to 1,000 meters produces a coefficient closer to the baseline (0.77). These movements in point estimates are plausibly driven by differences across the set of schools and neighborhoods that are dropped when the bandwidth is changed. For instance, denser parts of Montreal are unaffected by bandwidth restrictions since students living in these areas all live very close to a boundary. Large distances from boundaries are only observed in the suburbs—about half my sample of permanent residents live within 200 meters of their nearest boundary. Overall, most estimates shown in Figure A17 remain within short range of the baseline results.

A separate issue arises in the decomposition exercise: the RD estimates may reflect local average treatment effects for a different subpopulation than the one that identifies total exposure effects. To address concerns related to heterogeneous treatment effects, I use a nearest-neighbor matching algorithm to re-weight the sample of permanent residents so that their distribution of observables matches the one of the movers' subsample (Jann 2017, Abadie and Imbens 2011). School effects estimates for this reweighted sample are close to the baseline results (Table A12). Tests for locally constant effects around the cutoff (Dong and Lewbel 2015) also indicate that effects are plausibly of the same magnitude for students away from the cutoff (online Appendix D). In online Appendix D, I also verify that estimates of π are stable across boundaries that generate gaps in the quality of default options of different magnitudes, and calculate model-based estimates of π using the movers design—and therefore the mover sample—rather than the RD approach. All of the associated results are in line with the range of estimates of π reported in the main text.

B. Movers Estimates

Within-Family Exposure Effects.—A possible violation of the identifying assumption is that students with higher (lower) unobserved family inputs might move to higher (lower) educational attainment areas earlier. To account for any time-invariant family unobserved heterogeneity, I add household fixed effects to the model. Here, identification relies on age differences between siblings. Positive exposure effects then imply that the relationship between the change in neighborhood and school quality, on one hand, and the difference in educational outcomes among siblings, on the other hand, should vary proportionally with the age-difference of siblings.

Since siblings are not directly identified in the data, I match students using unique moves at a very fine level of geography. More precisely, I assume that two students who move from and to the exact same six-digit postal codes in the same year must belong to the same household.⁴⁵ In columns 1 and 2 of Table 5, I estimate

⁴⁵ Out of the original 100,929 students, this method identifies about 13,000 siblings attached to roughly 6,000 different households. Many household units are not consistent over time given the prevalence of step- and

TABLE 5—TOTAL EXPOSURE EFFECTS—SIBLINGS SUBSAMPLE

Sample:	Siblings only			
	(1)	(2)	(3)	(4)
<i>Measure of educational attainment</i>				
University enrollment	−0.0453 (0.0344)	−0.0571 (0.0339)	−0.0478 (0.0275)	−0.0504 (0.0274)
Secondary school diploma in five years	−0.0242 (0.0356)	−0.0434 (0.0344)	−0.0392 (0.0301)	−0.05 (0.0300)
Years of education	−0.0453 (0.0344)	−0.0629 (0.0339)	−0.0444 (0.0275)	−0.0486 (0.0274)
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		
Household fixed effects			X	X
Times in difficulty before moving		X		X
Observations	3,269	3,269	3,674	3,674

Notes: This table reports regression estimates of total exposure effects on educational attainment in a subsample of siblings. Estimates of the convergence rate β are estimated from a regression of movers outcomes on $\Delta \bar{y}_{od}$, the difference in outcomes between permanent residents of the destination and origin neighborhoods, interacted with age at move, as in Table 3, but the sample is now restricted to households in which siblings lived at the same address for at least 75 percent of the observed years. Students who moved from or to neighborhoods (FSAs) with less than ten permanent residents are always excluded. In columns 1 and 2, origin-by-destination fixed effects are included, whereas in columns 3 and 4 they are replaced with household fixed effects. Individual characteristics are gender, immigrant status, allophone status, born in Canada but outside Quebec, English spoken at home, day care use at baseline, “in difficulty” status at baseline, handicapped status. In columns 2 and 4, the model includes a set of dummies for each possible value of number of times a student is identified as being “in difficulty” prior to moving. Standard errors, shown in parentheses, are clustered at the household level.

the exposure model with origin-by-destination fixed effects on the subsample of siblings. Standard errors are considerably larger than in the main specification given the smaller sample size, but the point estimates are in line with the main results. In columns 3 and 4, I substitute family fixed effects for the origin-by-destination fixed effects to account for any time-invariant heterogeneity across families and still find convergence rates of about 4.5 percent. These results support the claim that the estimated exposure effects are not driven by differences in unobservable time-invariant family characteristics.

Balance of Observables.—I then test for balance of covariates to verify that variation in the interaction term is arguably random conditional on age-at-move and origin-by-destination fixed effects. Formally, I use individual characteristics as dependent variables in equation (4). Pei, Pischke, and Schwandt (2017) show that putting the covariates on the left-hand side is a more powerful test than gradually adding or removing these variables from the right-hand side of the main

blended-families. For instance, two students from different biological parents may have been living under the same roof only for a fraction of their lives. I therefore exclude household units for which the children have lived at a common postal code for less than 75 percent of the years for which I can observe them.

estimating equation, particularly if the individual characteristics are poor measures of the underlying confounders they are meant to account for (e.g., being “in difficulty” is certainly a noisy measure of academic abilities).

Results are shown in Table A13 in the online Appendix. The coefficients on immigrant status are marginally significant at the 5 percent level for some, but not all, outcomes. In Montreal, immigrants do obtain more postsecondary education than domestic students. It might also be the case that they tend to move later given that they may have less prior information about neighborhoods than native-born parents. Overall, most coefficients in the table are very small and statistically indistinguishable from zero. As a result, the baseline convergence rate is materially unchanged whether covariates are included or not, despite the fact that these covariates have nontrivial explanatory power with respect to educational attainment—e.g., for years of education, their inclusion increases the adjusted R^2 from 0.188 to 0.307.

Selection on Time-Varying Observables.—Another possible source of concern is that length of exposure to a given area mirrors exposure to different family circumstances. One may be worried that if a move is triggered by a change in marital status or income, age-specific unobserved parental inputs may also change in proportion with m_i . Unfortunately, my dataset includes very few time-varying individual characteristics. For instance, parental income and marital status of parents are not observed. To account for possible changes in family circumstances that coincide with a move, I instead control for differences in census tract characteristics between the areas in which student i resides after and before the move, as well as the interaction of these differences with age-at-move. Because census tracts are considerably smaller than FSAs, these controls vary within origin-by-destination cells, and so the main effects are identified. The characteristics I consider are median household income, average dwelling value, percentage low income, percentage of adults with some college education, and fraction of lone parent families. The inclusion of these variables accounts for changes in family circumstances that are correlated with changes in neighborhood attributes, as well as any sorting on the basis of these observable neighborhood characteristics. Similarly, Altonji and Mansfield (2014) control for group-level average individual characteristics—arguably the basis on which households sort into neighborhoods—to account for unobserved individual heterogeneity, and thereby obtain a lower bound on contextual effects. For example, a positive income shock may be associated with both a move to an area where property value is higher and an increase in parental inputs. For any unobserved variable to generate bias in this analysis, the confounding variable would have to generate variation orthogonal to changes in these neighborhood attributes. The inclusion of these neighborhood attributes likely absorbs part of the causal exposure effect of interest, and therefore overadjusts for changes in family circumstances.

Results are quite robust to the inclusion of these controls (online Appendix Table A14). In columns 1 through 5, I control for changes in one time-varying characteristic at a time. In all of these cases, the exposure effects remain stable around 4 to 4.5 percent. Among all considered variables, the local fraction of college-educated residents is the one that most affects the main exposure effects. Yet, even in this case, the exposure effects remain large ($\approx 4\%$). In column 6, I include all controls

simultaneously—exposure effects fall just under 4 percent and remain strongly statistically significant.

Event-Study.—Next, I investigate whether students who move to higher outcome areas exhibit different trends in learning difficulties prior to moving. The idea is that family circumstances plausibly directly affect the likelihood that a student struggles in school, hence changes in unobserved family inputs should be reflected in the probability of being identified in difficulty. I leverage year-to-year variation in $Diff_{i,od,t}$, the indicator of whether student i was in difficulty in year t , to create an index of relative learning difficulties that summarizes the way movers compare to permanent residents in their origin and destination $\sigma_{od(i,t)} = (Diff_{i,od,t} - \overline{Diff}_{o,t}) / (\overline{Diff}_{d,t} - \overline{Diff}_{o,t})$ where $\overline{Diff}_{n,t}$ is the fraction of permanent residents of FSA n that were in difficulty at time t (years since started grade 1). This index takes a value of zero if mover i 's difficulty status is the same as the average in her origin and a value of one if it is equal to the average in the destination. An increase in $\sigma_{od(i,t)}$ over time indicates that student i 's success in school (or lack thereof) converges toward that of permanent residents in the destination relative to the origin.

I investigate patterns in $\sigma_{od(i,t)}$ around the time of moves. For instance, a positive pre-trend would indicate that movers started converging toward permanent residents of the destination before they even moved. Such a pattern could arise if moves to certain areas occur as a result of gradual changes in family circumstances. For example, if divorces are preceded by an erosion of the quality of the parents' relationship *and* are disproportionately followed by moves to places with lower educational outcomes, my estimates of total exposure effects could be biased.

Online Appendix Figure A19 shows an event-study analysis in which $\sigma_{od(i,t)}$ is regressed on time dummies (relative to year of move). Observations are weighted by $(\overline{Diff}_{d,t} - \overline{Diff}_{o,t})^2$, as in Bronnenberg, Dubé, and Gentzkow (2012). A jump in $\sigma_{od(i,t)}$ occurs on impact, and students schooling difficulties then converge gradually toward the destination's average (panel A). Importantly, there is no discernible pre-trend—coefficients are stable prior to moving. These results are consistent with Aaronson (1998), who finds no systematic pattern between pre-move changes in family circumstances and the quality of the destination neighborhood quality: moves preceded by a divorce are just as likely to lead to a high or low outcome destination.

Because students move at different ages, the panel is not balanced. As a result, any pre- and post-move trends may be the result of changes in the sample's composition. In panel B, I follow Finkelstein, Gentzkow, and Williams (2016) and include student fixed effects to address this issue. While the post-move trend now disappears, the jump at the time of the move remains significant. One concern is that this sudden jump is the product of a sharp and sudden change in family inputs. In panels C and D, I therefore distinguish between students who did and did not switch school the year they moved. Students who did not change school at the time of the move show no jump in $\sigma_{od(i,t)}$, suggesting that the break reflects a change in the schooling environment (e.g., differences in schools' propensity to flag marginal students) rather than a change in family inputs. Overall, I find no evidence of pre-trends in schooling difficulties.

C. Decomposition

Movers' School Attendance.—The main decomposition approach projects differences in the quality of schools attended by permanent residents onto movers. One concern is that measures of school quality estimated on a sample of nonmovers may not accurately capture the true change in school quality experienced by movers. In this subsection, I instead directly account for movers' school attendance in the baseline total exposure effect model. The estimating equation becomes

$$(16) \quad y_{icmod} = \beta(m_i \times \Delta \bar{y}_{od}) + \alpha_{s(0)} + \alpha_{s(A)} + \gamma X_{icod} + \alpha_{od} + \alpha_m + \alpha_c + \varepsilon_{icmod},$$

where $\alpha_{s(0)}$ and $\alpha_{s(A)}$ are sets of fixed effects for schools attended at baseline and at age 15, respectively. To account for variation in length of exposure to different schools that may be correlated with neighborhood exposure, the school fixed effects are allowed to vary linearly with age-at-move, i.e., $\alpha_{s(a)} = \alpha_{s(a)}^0 + \alpha_{s(a)}^1 \times m_i$, which is equivalent to allowing age-at-move effects to have a different slope in each school.

The results are presented in Table 6. Benchmark estimates of total exposure effects are reproduced in column 1. In column 2, fixed effects for schools attended at the beginning (the “origin” school) and at the end (the “destination” school) of the exposure period, as well as interactions with age-at-move, are added. Net of movers' school attendance, the annual exposure effects shrink substantially to 1.1 percent for university enrollment, 1.2 percent for completing secondary school on time, and 1.8 percent for years of education. As a further robustness check, in column 3 the school fixed effects are interacted with the actual number of years spent in the associated schools instead of with age-at-move.

To get a sense of how these results compare to the main decomposition results, I report corresponding estimates of $\beta^{non-school}$ in column 4. Because the inclusion of school fixed effects in equation (16) absorbs any variation associated with school attendance (i.e., it may account for more than just the causal component), I report estimates of $\beta^{non-school}$ for which π is set to 1 to make the two sets of estimates comparable. Looking across columns 2 and 4, one finds the point estimates to be strikingly close (e.g., 1.1 percent versus 1.0 percent for university enrollment), suggesting that $\Delta \Omega_{od}$ does a good job of representing the effective change in school quality faced by movers.

Sampling Error.—Movers design estimates likely suffer from attenuation bias due to sampling error in the school and neighborhood fixed effects estimates. I consider two approaches to addressing this issue. First, I use empirical Bayes techniques (Chandra et al. 2016; Best, Hjort, and Szakonyi 2017; Kane and Staiger 2008) to shrink toward zero the school and neighborhood fixed effects and use these shrunk estimates as regressors in equations (4), (13), and (14). Second, I use a split-sample instrumental variable approach in which I randomly split the sample of permanent residents in half and obtain two sets of estimates of school and neighborhood quality. I then instrument for $\Delta \bar{y}_{od,1}$ with $\Delta \bar{y}_{od,2}$, $\Delta \Omega_{od,1}$ with $\Delta \Omega_{od,2}$, and $\Delta \bar{y}_{od,1}^{-s}$ with $\Delta \bar{y}_{od,2}^{-s}$. This approach addresses the possibility that the signal-to-noise

TABLE 6—EXPOSURE EFFECTS NET OF MOVERS' SCHOOL ATTENDANCE

	Convergence rate			Decomposition: $\beta^{non-school}$ for $\pi = 1$
	(1)	(2)	(3)	(4)
<i>Measure of educational attainment</i>				
University enrollment	−0.0424 (0.0090)	−0.0111 (0.0101)	−0.0123 (0.0089)	−0.0103 (0.0038)
Secondary school diploma in five years	−0.0421 (0.0088)	−0.0117 (0.0093)	−0.0075 (0.0081)	−0.0116 (0.0039)
Years of education	−0.0488 (0.0088)	−0.0178 (0.0079)	−0.0181 (0.0075)	−0.0138 (0.0034)
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
School fixed effects				
(o) School at baseline		X	X	
(o) School at baseline \times age-at-move (linear)		X		
(o) School at baseline \times years-exposure			X	
(d) School at age 15		X	X	
(d) School at age 15 \times age-at-move (linear)		X		
(d) School at age 15 \times years-exposure			X	
Observations	24,316	24,244	24,244	24,316

Notes: This table reports regression estimates of total exposure effects on educational attainment controlling for movers' school attendance. Estimates of the convergence rate β are estimated from a regression of movers outcomes on $\Delta \bar{y}_{od}$, the difference in outcomes between permanent residents of the destination and origin neighborhoods, interacted with age at move. The sample includes all movers, but students who moved from or to FSAs with less than ten permanent residents are excluded. Column 1 replicates total exposure effects reported in Table 3. In columns 2 and 3, school fixed effects are added to the model. These fixed effects are linearly interacted with age-at-move in column 2, and with years spent in that school in column 3. The "origin" school is the primary school attended at baseline, and the "destination" school is the secondary school attended at age 15. Standard errors, shown in parentheses, are clustered at the destination neighborhood level. Column 4 reports estimates of the non-school component of exposure effects $\beta^{non-school}$ obtained from a decomposition analysis and presented in Table 4, column 1.

ratio may be greater in the movers design—which relies on residual variation net of origin-by-destination fixed effects—than the reliability ratios reported in Table 1.

The results are shown in online Appendix Table A15. The first row of panel A shows convergence rates β estimated using shrunk estimates of permanent residents' mean outcomes \bar{y}_n^{EB} in equation (4). These rates are slightly larger than the ones presented in Table 3, varying between 4.5 percent and 5 percent, which imply that my main estimates suffer from a small attenuation bias. The second row reports "two-step" convergence rates obtained using shrunk estimates of school and neighborhood fixed effects, $\Omega_{s(i,n)}^{EB}$ and Λ_n^{EB} . The two rates differ because the sum of two shrunk estimates is not the same as shrunk estimates of the sum. While $\bar{y}_n = \Lambda_n + \bar{\Omega}_n$ holds true by construction, the equality does not hold for shrunk estimates: $\bar{y}_n^{EB} \neq \Lambda_n^{EB} + \bar{\Omega}_n^{EB}$. For all three outcomes, adjusting for measurement error does not alter the conclusion that school effects account for most of the benefits of moving to a high educational attainment area. Panel B presents the results for the split-sample IV approach. The total rates β are very close to those in panel A, again suggestive of a small attenuation bias. School shares adjusted for π here vary between 64 percent and 72 percent, a magnitude slightly greater than the main results.

VII. Conclusion

Establishing the relative magnitude of school and neighborhood exposure effects is crucial on a policy level to inform the development of community-wide versus in-school intervention programs. In this paper, I have estimated the long-term impact of growing up in better neighborhoods and attending better schools on educational attainment. I then implemented a decomposition approach which, under several strong assumptions, provides an estimate of the fraction of the effect of growing up in a given area that is due to schools.

Using a spatial regression discontinuity design based on unique institutional features of Quebec's education system, I find that the primary and secondary schools children attend have large effects on their educational attainment. More precisely, immediate neighbors living on opposite sides of a French primary school boundary at age 6 exhibit significantly different propensities to enroll into university more than 10 years later. My second set of results demonstrates that children who move to a high educational attainment neighborhood at a young age benefit substantially from this change. I find strong evidence in favor of linear exposure effects, in line with US findings (Chetty and Hendren 2018a, Chetty et al. 2018). My estimates suggest that movers' educational attainment converges to that of permanent residents at an annual rate of 4.5 percent.

Then, I provide a new method to quantify mechanisms in mover designs and find that schools are the main driver of total childhood exposure effects. Results that take into account the endogeneity of school quality indicate that between 50 percent and 70 percent of the effects of moving to a given location are due to differences in school quality. These findings strongly corroborate earlier conclusions made on the relative importance of schools and neighborhoods (Dobbie and Fryer 2015, Fryer and Katz 2013, Oreopoulos 2012) and show that spatial inequalities in long-term educational attainment are partly rooted in the quality of schools children attend. They notably suggest that school reforms or interventions might be more effective than community programs or relocation policies in raising educational attainment.

Some limitations of this study must be acknowledged. The decomposition approach hinges on a constant treatment effects assumption. If exposure effects are heterogeneous, my estimates of place effects may suffer from an attenuation bias. Also, the way school quality is measured conflates many different inputs, such as peer, teacher and school principal quality. While the magnitude of the estimated exposure effects, and the main conclusion of this paper more broadly, may reflect a social reality unique to Montreal, I believe the results are very informative for other contexts as well. For instance, because of Quebec's open enrollment policy and the unusual availability of private school options in Montreal, the link between school attendance and residence is relatively loose. Hence, in jurisdictions where schools and neighborhoods are tightly linked, one may expect schools to contribute even more to spatial inequalities in educational attainment than the results in this paper suggest. An important question for further research is whether these results generalize to other noneducational outcomes, such as earnings and criminal behavior.

REFERENCES

- Aaronson, Daniel. 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33 (4): 915–46.
- Abadie, Alberto, and Guido W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business and Economic Statistics* 29 (1): 1–11.
- Abdulkadiroğlu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics* 126 (2): 699–748.
- Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica* 82 (1): 137–96.
- Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2017. "Do Parents Value School Effectiveness?" NBER Working Papers 23912.
- Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.
- Agrawal, Mohit, Joseph G. Altonji, and Richard K. Mansfield. 2018. "Quantifying Family, School, and Location Effects in the Presence of Complementarities and Sorting." NBER Working Paper 25167.
- Altonji, Joseph G., and Richard K. Mansfield. 2014. "Group-Average Observables as Controls for Sorting on Unobservables When Estimating Group Treatment Effects: The Case of School and Neighborhood Effects." NBER Working Papers 20781.
- Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *Quarterly Journal of Economics* 132 (2): 871–919.
- Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (4): 979–1014.
- Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2016. "School Quality and the Gender Gap in Educational Achievement." *American Economic Review* 106 (5): 289–95.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.
- Best, Michael Carlos, Jonas Hjort, and David Szakonyi. 2017. "Individuals and Organizations as Sources of State Effectiveness, and Consequences for Policy." NBER Working Paper 23350.
- Billings, Stephen B., David J. Deming, and Jonah Rockoff. 2014. "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg." *Quarterly Journal of Economics* 129 (1): 435–76.
- Billings, Stephen B., David J. Deming, and Stephen L. Ross. 2016. "Partners in Crime: Schools, Neighborhoods and the Formation of Criminal Networks." NBER Working Paper 21962.
- Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (2): 577–99.
- Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell. 2010. "The Evolution of the Returns to Human Capital in Canada, 1980–2005." *Canadian Public Policy* 36 (1): 63–89.
- Bronnenberg, Bart J., Jean-Pierre H. Dubé, and Matthew Gentzkow. 2012. "The Evolution of Brand Preferences: Evidence from Consumer Migration." *American Economic Review* 102 (6): 2472–508.
- Burdick-Will, Julia, Jens Ludwig, Stephen W. Raudenbush, Robert J. Sampson, Lisa Sanbonmatsu, and Patrick Sharkey. 2011. "Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-experimental, and Observational Comparison." In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Greg J. Duncan and Richard J. Murnane, 255–76. New York: Russell Sage Foundation.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–2326.
- Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127–60.
- Card, David, Ciprian Domnisoru, and Lowell Taylor. 2018. "The Intergenerational Transmission of Human Capital: Evidence from the Golden Era of Upward Mobility." NBER Working Paper 25000.
- Card, David, Martin D. Dooley, and A. Abigail Payne. 2010. "School Competition and Efficiency with Publicly Funded Catholic Schools." *American Economic Journal: Applied Economics* 2 (4): 150–76.
- Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128 (3): 967–1015.

- Card, David, and Jesse Rothstein.** 2007. "Racial Segregation and the Black-White Test Score Gap." *Journal of Public Economics* 91 (11–12): 2158–84.
- Carlson, Deven, and Joshua M. Cowen.** 2015. "Student Neighborhoods, Schools, and Test Score Growth: Evidence from Milwaukee, Wisconsin." *Sociology of Education* 88 (1): 38–55.
- Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson.** 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review* 106 (8): 2110–44.
- Chetty, Raj.** 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review* 105 (5): 1–33.
- 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." NBER Working Paper 25147.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff.** 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79.
- Chetty, Raj, John N. Friedman, and Emmanuel Saez.** 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.
- Chetty, Raj, and Nathaniel Hendren.** 2018a. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *Quarterly Journal of Economics* 133 (3): 1107–62.
- Chetty, Raj, and Nathaniel Hendren.** 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *Quarterly Journal of Economics* 133: 1163–1228.
- Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz.** 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.
- Chyn, Eric.** 2018. "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Children." *American Economic Review* 108 (10): 3028–56.
- Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt.** 2006. "The Effect of School Choice on Participants: Evidence from Randomized Lotteries." *Econometrica* 74 (5): 1191–1230.
- Damm, Anna Piil, and Christian Dustmann.** 2014. "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?" *American Economic Review* 104 (6): 1806–32.
- Deming, David J.** 2014. "Using School Choice Lotteries to Test Measures of School Effectiveness." *American Economic Review* 104 (5): 406–11.
- Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger.** 2014. "School Choice, School Quality, and Postsecondary Attainment." *American Economic Review* 104 (3): 991–1013.
- Dobbie, Will, and Roland G. Fryer, Jr.** 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3 (3): 158–87.
- Dobbie, Will, and Roland G. Fryer, Jr.** 2013. "Getting beneath the Veil of Effective Schools: Evidence from New York City." *American Economic Journal: Applied Economics* 5 (4): 28–60.
- Dobbie, Will, and Roland G. Fryer, Jr.** 2015. "The Medium-Term Impacts of High-Achieving Charter Schools." *Journal of Political Economy* 123 (5): 985–1037.
- 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): 1081–92.
- Duhaime-Ross, Alix.** 2015. "Three Essays in the Economics of Education: Evidence from Canadian Policies." PhD diss. University of British Columbia.
- Fack, Gabrielle, and Julien Grenet.** 2010. "When Do Better Schools Raise Housing Prices? Evidence from Paris Public and Private Schools." *Journal of Public Economics* 94 (1–2): 59–77.
- Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams.** 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics* 131 (4): 1681–1726.
- Fryer, Roland G., Jr., and Lawrence F. Katz.** 2013. "Achieving Escape Velocity: Neighborhood and School Interventions to Reduce Persistent Inequality." *American Economic Review* 103 (3): 232–37.
- Gannon, Maire.** 2001. "Crime Comparisons between Canada and the United States." *Juristat—Canadian Centre for Justice Statistics* 21 (11): 1–13.
- Gibbons, Stephen, Stephen Machin, and Olmo Silva.** 2013. "Valuing School Quality Using Boundary Discontinuities." *Journal of Urban Economics* 75: 15–28.
- Gibbons, Stephen, Olmo Silva, and Felix Weinhardt.** 2013. "Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England." *Economic Journal* 123 (571): 831–74.

- Gould, Eric D., Victor Lavy, and M. Daniele Paserman.** 2004. "Immigrating to Opportunity: Estimating the Effect of School Quality Using a Natural Experiment on Ethiopians in Israel." *Quarterly Journal of Economics* 119 (2): 489–526.
- Gould, Eric D., Victor Lavy, and M. Daniele Paserman.** 2011. "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes." *Review of Economic Studies* 78 (3): 938–73.
- Goux, Dominique, and Eric Maurin.** 2007. "Close Neighbours Matter: Neighbourhood Effects on Early Performance at School." *Economic Journal* 117 (523): 1193–1215.
- Hanushek, Eric A.** 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24 (3): 1141–77.
- Heckman, James J., John Eric Humphries, and Gregory Veramendi.** 2017. "The Non-market Benefits of Education and Ability." NBER Working Papers 23896.
- Heckman, James J., Jora Stixrud, and Sergio Urzua.** 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–82.
- Houtenville, Andrew J., and Karen Smith Conway.** 2008. "Parental Effort, School Resources, and Student Achievement." *Journal of Human Resources* 43 (2): 437–53.
- Hoxby, Caroline M.** 2015. *Computing the Value-Added of American Postsecondary Institutions*. Washington, DC: Internal Revenue Service (IRS), US Department of Treasury.
- Imberman, Scott A., and Michael F. Lovenheim.** 2016. "Does the Market Value Value-Added? Evidence from Housing Prices after a Public Release of School and Teacher Value-Added." *Journal of Urban Economics* 91: 104–21.
- Jackson, C. Kirabo.** 2010. "Do Students Benefit from Attending Better Schools?: Evidence from Rule-Based Student Assignments in Trinidad and Tobago." *Economic Journal* 120 (549): 1399–1429.
- Jackson, C. Kirabo.** 2016. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-test Score Outcomes." NBER Working Paper 22226.
- Jacob, Brian A.** 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94 (1): 233–58.
- Jann, Ben.** 2017. "KMATCH: Stata Module for Multivariate-Distance and Propensity-Score Matching." Statistical Software Components, Boston College Department of Economics. <https://boris.unibe.ch/113482/> (accessed February 23, 2021).
- Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger.** 2006. "School Quality, Neighborhoods, and Housing Prices." *American Law and Economics Review* 8 (2): 183–212.
- Kane, Thomas J., and Douglas O. Staiger.** 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." NBER Working Paper 14607.
- Katz, Lawrence F.** 2015. "Reducing Inequality: Neighborhood and School Interventions." *Focus* 31 (2): 12–17.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.
- Laliberté, Jean-William.** 2021. "Replication data for: Long-Term Contextual Effects in Education: Schools and Neighborhoods." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E118906V1>.
- Lapierre, David, Pierre Lefebvre, and Philip Merrigan.** 2016. "Long Term Educational Attainment of Private High School Students in Québec: Estimates of Treatment Effects from Longitudinal Data." University of Quebec in Montreal's School of Management, Research Group on Human Capital Working Paper 16-02.
- Lavecchia, Adam M., Heidi Liu, and Philip Oreopoulos.** 2014. "Behavioral Economics of Education: Progress and Possibilities." NBER Working Paper 20609.
- Lee, David S., and Thomas Lemieux.** 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.
- Lefebvre, Pierre, Philip Merrigan, and Matthieu Verstraete.** 2011. "Public Subsidies to Private Schools Do Make a Difference for Achievement in Mathematics: Longitudinal Evidence from Canada." *Economics of Education Review* 30 (1): 79–98.
- Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu.** 2013. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity." *American Economic Review* 103 (3): 226–31.
- McCrary, Justin.** 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714.

- Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88 (9–10): 1667–95.
- Ministère de l'Éducation, de l'Enseignement supérieur et de la Recherche. 2016. "Base de données administratives [Database]."
- Molitor, David. 2018. "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration." *American Economic Journal: Economic Policy* 10 (1): 326–56.
- Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* 121 (1–2): 175–212.
- Oreopoulos, Philip. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood." *Quarterly Journal of Economics* 118 (4): 1533–75.
- Oreopoulos, Philip. 2008. "Neighborhood Effects in Canada: A Critique." *Canadian Public Policy—Analyse de Politiques* 34 (2): 237–58.
- Oreopoulos, Philip. 2012. "Moving Neighborhoods versus Reforming Schools: A Canadian's Perspective." *Cityscape* 14 (2): 207–12.
- Oreopoulos, Philip, and Uros Petronijevic. 2013. "Making College Worth It: A Review of Research on the Returns to Higher Education." NBER Working Paper 19053.
- Oreopoulos, Philip, and Kjell G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives* 25 (1): 159–84.
- Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt. 2017. "Poorly Measured Confounders Are More Useful on the Left Than on the Right." NBER Working Paper 23232.
- Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review* 103 (4): 1289–1324.
- Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–58.
- Rothstein, Jesse M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review* 96 (4): 1333–50.
- Rothstein, Jesse. 2017. "Inequality of Educational Opportunity? Schools as Mediators of the Intergenerational Transmission of Income." https://eml.berkeley.edu/~jrothst/workingpapers/rothstein_mobility_april2017.pdf.
- Sharkey, Patrick, and Jacob W. Faber. 2014. "Where, When, Why, and for Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects." *Annual Review of Sociology* 40: 559–79.
- Sykes, Brooke, and Sako Musterd. 2011. "Examining Neighbourhood and School Effects Simultaneously: What Does the Dutch Evidence Show?" *Urban Studies* 48 (7): 1307–31.
- Todd, Petra E., and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal* 113 (485): F3–F33.
- Weinhardt, Felix. 2014. "Social Housing, Neighborhood Quality and Student Performance." *Journal of Urban Economics* 82: 12–31.
- Wodtke, Geoffrey T., and Matthew Parbst. 2017. "Neighborhoods, Schools, and Academic Achievement: A Formal Mediation Analysis of Contextual Effects on Reading and Mathematics Abilities." *Demography* 54 (5): 1653–76.
- Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics* 32 (4): 711–54.