

NBER WORKING PAPER SERIES

WHAT IS A GOOD SCHOOL, AND CAN PARENTS TELL? EVIDENCE ON THE
MULTIDIMENSIONALITY OF SCHOOL OUTPUT

Diether Beuermann
C. Kirabo Jackson
Laia Navarro-Sola
Francisco Pardo

Working Paper 25342
<http://www.nber.org/papers/w25342>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2018, Revised October 2021

We are deeply grateful to Sabine Rieble-Aubourg and Dana King from the Inter-American Development Bank for their invaluable support in establishing the necessary contacts to assemble the data used in the study. We are indebted to Chief Education Officer Harrilal Seecharan of the Trinidad and Tobago Ministry of Education (TTMOE) for his continuous support. We also thank Ria Boafó, Lisa Henry-David, Shalini Maharaj, Brenda Moore, and Peter Smith of the TTMOE for facilitating access to the educational data needed for the study, their assistance, and their generosity. We thank Registrar General of Trinidad and Tobago Karen Bridgewater for kindly granting us access to the national birth records; Amos Sylvester from the Crime and Problem Analysis Branch of the Trinidad and Tobago Police Service for facilitating access to arrests records; and Executive Director Niala Persad of the National Insurance Board of Trinidad and Tobago, as well as Andy Edwards, Arlene Grant, Feyaad Khan and Bernard Smith for their support and generosity while working in their facilities to match employment records while maintaining individual confidentiality. Tatiana Zárate and Diego Zúñiga provided excellent research assistance. This paper benefited from comments by seminar participants at the NBER, MIT, UChicago, LACEA, Zurich, Lund, and Wharton. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Diether Beuermann, C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output

Diether Beuermann, C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo

NBER Working Paper No. 25342

December 2018, Revised October 2021

JEL No. I2,J01,J38

ABSTRACT

To explore whether schools' causal impacts on test scores measure their overall impact on students, we exploit plausibly exogenous school assignments and data from Trinidad and Tobago to estimate the causal impacts of individual schools on several outcomes. Schools' impacts on high-stakes tests are weakly related to impacts on important outcomes such as arrests, dropout, teen motherhood, and formal labor-market participation. To examine if parents' school preferences are related to these causal impacts, we link them to parents' ranked lists of schools and employ discrete-choice models to infer preferences for schools. Parents choose schools that improve high-stakes tests even conditional on peer quality and average outcomes. Parents also choose schools that reduce criminality and teen motherhood, and increase labor market participation. School choices among parents of low-achieving students are relatively more strongly related to schools' impacts on non-test-score outcomes, while the opposite is true for parents of high-achieving students. These results suggest that evaluations based solely on test scores may be misleading about the benefits of school choice (particularly for low-achieving students), and education interventions more broadly.

Diether Beuermann
Inter-American Development Bank
1300 New York Ave, NW
Washington, DC 20577
dietherbe@iadb.org

Laia Navarro-Sola
Institute for International Economic Studies
Stockholm University
Stockholm SE-106 91
Sweden
laia.navarrosola@iies.su.se

C. Kirabo Jackson
Northwestern University
School of Education and Social Policy
Annenberg Hall, #204
2120 Campus Dr.
Evanston, IL 60208
and NBER
kirabo-jackson@northwestern.edu

Francisco Pardo
Inter-American Development Bank
pardopajuelofrancisco@gmail.com

I Introduction

Is a school’s causal impact on test scores a good measure of its overall impact on students? Do parents value schools that improve high-stakes standardized tests? Do parents value school impacts on outcomes other than high-stakes tests? To shed light on these issues, we use administrative data from many sources covering the full population of Trinidad and Tobago. To address the first question, we estimate individual schools’ causal impacts on high-stakes test scores, low-stakes test scores, dropout, teen motherhood, teen arrests, and labor market participation. Using the relationship between these estimates, we examine whether school output is multidimensional (such that test score impacts may not be related to school impacts on other dimensions). To address the second and third questions, we link our estimated impacts to parents’ school rankings and explore whether parents choose schools with causal impacts on these multiple outcomes – providing the first exploration into whether parents may value school causal impacts on non-academic outcomes.

The motivations for this paper are twofold. First, we aim to better understand the multidimensional nature of individual schools’ output. Researchers, practitioners, and policy-makers often rely on school’s performance on standardized tests as a measure of quality. However, because educational output may be multidimensional (Hanushek 1971; Heckman et al. 2006; Kautz et al. 2017; Jackson 2018; Jackson et al. forthcoming), schools that improve important longer-run outcomes (such as crime, college-going, and earnings) may have little impact on test scores. As such, policies that focus primarily on test score impacts to make decisions (such as school closures, performance pay, accountability, etc.) may not necessarily improve longer-run outcomes that policymakers and parents value. To assess the importance of this, one must understand the joint distribution of individual school’s impacts across several different outcomes.¹ However, to date, only four studies examine the causal impact of individual schools on different outcomes (Abdulkadiroğlu et al. 2014; Dobbie and Fryer 2015; Angrist et al. 2016; Place and Gleason 2019). To rely on *causal* impacts, these studies focus on a small number of oversubscribed schools that admit students using enrollment exams or randomized lotteries.² While focusing on oversubscribed schools

¹We now know that certain *groups of schools* that raise test scores may not improve other outcomes and vice versa. For example, Deming (2011) finds that winning a school choice lottery can reduce crime with little impact on test scores, Deming et al. (2014) find that school choice lotteries improve test scores and educational attainment (only for girls). Beuermann and Jackson (2020) find that attending a preferred school in Barbados improves long run outcomes but not test scores. Also, Booker et al. (2011) find that charter school attendance impacts on test score do not correlate with their impacts on college outcome. All these studies examine groups of schools rather than individual school impacts – precluding an analysis of the multidimensional nature of educational output by schools.

²Place and Gleason (2019) and Angrist et al. (2016) examine 31 and 26 oversubscribed charter schools, respectively. Abdulkadiroğlu et al. (2014) examine 6 elite selective enrollment schools, and Dobbie and Fryer (2015) examine a single charter school. Dobbie and Fryer (2020) examine impacts of 57 charter schools in Texas. However, they do not have exogenous variation so that the estimated school impacts may not capture causal impacts. In their words “we assume unobserved determinants of students’ labor market outcomes are orthogonal to our school value-added measures....our estimates should be interpreted with this strong identifying assumption in mind.”

overcomes selection biases, these studies examine a small number of schools that are necessarily non-representative – limiting generalizability. Moreover, these studies examine individual schools’ impacts on test scores and related educational outcomes (such as college going) but *do not* relate schools’ test score impacts to a broad set of nonacademic outcomes. As such, no studies have used a quasi-experimental design to identify individual school’s causal impacts across a representative group of schools and on a broad array of academic and non-academic outcomes simultaneously – which is necessary to rigorously explore the multidimensional nature of school output. To help fill this space, we rely on plausibly exogenous variation to uncover the causal impact of attending 132 individual public secondary schools in Trinidad and Tobago (98.5 percent of all public secondary schools) on a wide array of academic and non-academic short- and longer-run outcomes.³

The second motivation for our work is to better understand parental preferences for schools. In theory, by aligning schools’ incentives with parents’ preferences, school choice policies may increase efficiency in educational production (Friedman 1955; Chubb and Moe 1990). However, if parents cannot discern school causal impacts, school choice policies will do little to increase education production or improve human capital. Indeed, there is a growing literature showing that parental preferences for schools are not systemically related to school impacts on test scores (MacLeod and Urquiola 2019; Beuermann and Jackson 2020). The few studies that directly examine preferences for school causal impacts conclude that parents may not value school impacts *per se* (Rothstein 2006, Abdulkadiroğlu et al. 2020) – casting doubt on the likely efficacy of school choice. However, there are two reasons that this may not be the final word on this issue; (1) Parents may value schools that improve outcomes largely unrelated to test-score impacts. If so, school choice may improve outcomes valued by parents, but that are not well observed by the econometrician – leading to wrong inferences about parental preferences and the benefits of school choice. (2) The only study to link secondary schools’ causal impacts to parents’ school choices (Abdulkadiroğlu et al. 2020) does so in New York City, the largest public school district in the United States. Because New York City is a relatively low-information setting where “overchoice” (Schwartz 2004; Iyengar and Kamenica 2010) may lead to sub-optimal decision making, it is unclear that their results would generalize to smaller and higher-information settings. Trinidad and Tobago provides such an alternate setting. By linking our school impacts on a broad set of outcomes to parents’ rankings of schools, we provide the first examination of the extent to which parents tend to choose schools that causally improve test scores and *also* key nonacademic outcomes. Our study, therefore, is the first to explore the relationship between school preferences and schools’ causal impacts in a high-information setting where problems of overchoice are more limited.

We use data on all applicants to public secondary schools in Trinidad and Tobago between 1995 and 2012. These data contain students’ identifying information, scores on the Secondary Entrance

³This is about the same number of public secondary schools in Chicago, and more than in the state of Vermont.

Assessment (taken at age 11 at the end of 5th grade), and a ranked list of secondary schools the student wished to attend. We link these data (at the student level) to scores on low-stakes national exams taken three years later, high-stakes secondary school completion exams five years later, and a national tertiary certification exam taken seven years later. We also link these student records to official police arrest records, birth registry data, and retirement contribution fund data. We track individual students over time through 33 years of age across a host of different types of outcomes.

To estimate the causal effects of attending individual schools, we rely on the fact that the Ministry of Education assigns most students to schools using a deferred acceptance algorithm (Gale and Shapley, 1962). School assignments are based on both student choices and student test scores. Conditional on the information used in the assignment process, the algorithm-based assigned school is unrelated to both observed and unobserved student characteristics (Jackson, 2010). Exploiting this fact, for each secondary school, we use the exogenous school assignment to that school as an instrument for *attending* that school (relative to a well-defined set of counterfactual schools). We implement several tests to support a causal interpretation of our estimates, and show that our relative school effects are approximately homogeneous. As such, differences in our effect estimates across any two schools reflect the relative improvement in outcomes most children can expect from attending one of these schools compared to the other— akin to others in the school effects literature (e.g., Deutsch et al. 2020, Abdulkadiroğlu et al. 2014; Dobbie and Fryer 2015; Angrist et al. 2016; Dobbie and Fryer 2020; Place and Gleason 2019).

To infer parental preferences for schools, we rely on the fact that a ranked list of four secondary schools is submitted as part of the secondary-school application process. Under Deferred Acceptance algorithms with unlimited choices, it is rational to list schools in order of true preference. Accordingly, standard rank-ordered logit models are identified on the assumption that the top choice is preferred to all other schools, that the second is second-most preferred and so on. Therefore, these models identify preference parameters under rational behaviour. However, under Deferred Acceptance algorithms with limited choices, *as in our setting*, (a) not all school rankings can be observed, and (b) applicants may be strategic by accounting for the likelihood of admission when making choices (Chade and Smith, 2006) – so that the “truthful revealing” identifying assumption may not apply.⁴ To account for this, we use estimators that assume that behaviours *may be* strategic, and therefore identify preferences under the proposed strategic behaviours— which are suitable for our context. Intuitively, because the nature of strategic choices is a function of admission probabilities, and we can obtain estimates of admission probabilities using historical data, we can model how admission probabilities influence choices and uncover true preferences for

⁴Researchers have addressed this by making some additional assumptions. Abdulkadiroğlu et al. (2020) assume that parents in New York City do not choose schools outside of their borough because such choices are uncommon. Also, to account for strategic choices both Abdulkadiroğlu et al. (2020) and Hastings et al. (2009) appeal to patterns in the data to justify the assumption that choices made are not strategic.

schools. Specifically, we implement a modified multinomial logit model (McFadden, 1973) modelling choices across *all* schools (making no assumption about what schools are in the choice set) and explicitly account for possible strategic behaviours and other constraints (such as proximity) that may cause an individual to not list her most preferred school as her top choice. Showing that our findings are not an artifact of the chosen methodology, our main findings are similar in models that do not account for strategic behaviors (as has been done by other researchers).

Schools have meaningful causal impacts on many outcomes. The standard deviation of school impacts on low-stakes and high-stakes exams is about 0.45σ . The standard deviation of school impacts are about 9 percentage points for dropout, 4 percentage points for teen arrests, 17 percentage points for teen births, and 7 percentage points for formal labor-market participation. We next test for whether school impacts on test scores capture effects on other outcomes. After accounting for estimation errors, the correlations between school impacts on high-stakes tests and other outcomes are modest. For example, the correlation between impacts on high-stakes exams and non-dropout is 0.12, and that between impacts on high-stakes tests and being formally employed is 0.15. We show that these low correlations are not due to high-achieving students being more responsive to school impacts on academic outcomes and attending one set of schools while low-achieving students being responsive to school impacts on nonacademic outcomes and attending another set of schools. Rather, *even among a homogeneous set of students*, schools that improve high-stakes test scores are often not those that improve broader adult well-being (which parents may value).

Next we analyze parental school choices to explore if parents' school preferences are related to schools' causal impacts on various outcomes. We replicate results of existing studies. Parents assign higher rankings to more proximate schools, higher-performing schools, and those with higher-achieving peers (Burgess et al. 2015; Abdulkadiroğlu et al. 2020). We also present several novel results. Conditional on peer quality, proximity, and school-level average outcomes; parents of higher-achieving children choose schools with larger positive causal impacts on high-stakes exams. This pattern cannot be driven by treatment heterogeneity because school impacts are largely homogeneous. These findings (from a high-information, modest-sized market) differ from Abdulkadiroğlu et al. (2020) who find that conditional on peer quality, parental preferences are unrelated to schools' test-score impacts in a large low-information setting – suggesting that information and/or overchoice may play a role.

Looking to nonacademic outcomes, we find robust evidence that parents prefer schools that reduce arrests, reduce teen births, and increase formal labor-market participation. However, there are key differences by student type. High-achieving students' choices are relatively more strongly related to schools' impacts on high-stakes exams than impacts on these nonacademic outcomes, while the choices of low-achieving students' are relatively more strongly related to schools' impacts on these non-academic outcomes than those on high-stakes exams. Because school impacts on these

outcomes are largely the same for students throughout the incoming test-score distribution, we can rule out that our key results are driven by test-score impacts leading to more improved outcomes for high-achieving children while labor market, teen birth, or crime impacts leading to more improved outcomes for lower-achieving students. Because school impacts on these outcomes are largely the same for students throughout the incoming test-score distribution, we can rule out that our key results are driven by test-score impacts leading to more improved outcomes for high-achieving children while labor market, teen birth, or crime impacts leading to more improved outcomes for lower-achieving students. Because schools that improve test scores may not reduce teen motherhood, crime or improve labor market participation, these results have important implications for our understanding of parental preferences for schools – particularly those of unprivileged populations.

We build on the school quality literature by presenting the first analysis of the relationships between schools' *causal* impacts on several academic and non-academic outcomes – providing direct evidence of the multidimensionality of school output.⁵ Our findings have important policy implications because test-score impacts are increasingly used for policy decisions. We also contribute to the work on parental preferences by providing the first study of parental choices based on school impacts on non-academic outcomes such as fertility, crime, and labor-market participation. We show that parents may have strong preferences for schools that reduce crime, reduce teen births, and increase labor market participation – impacts that are only weakly correlated with impacts on test scores. If this pattern holds in other settings, it could explain why researchers have found a weak link between parental preferences for schools and schools' test score impacts. As such, our results suggest that existing evaluations of school choice based solely on test-score impacts (without regard for schools' nonacademic output) may be very misleading about their welfare effects.

The remainder of this paper is as follows; Section II describes the Trinidad and Tobago context and discusses the data. Section III presents our empirical strategy for estimating school causal impacts. Section IV presents the magnitudes of the estimated school impacts and explores the potential multidimensionality of school output. Section V discusses our choice models, and presents estimates of the correlates of parental preferences. Section VI concludes.

II The Trinidad and Tobago Context and Data

The Trinidad and Tobago education system evolved from the English system. At the end of primary school (after grade 5, around 11 years old), parents register their children to take the Secondary Entrance Assessment (SEA) and provide a list of four ranked secondary school choices to the Ministry of Education (MOE). The SEA is comprised of five subjects that all students take:

⁵Dobbie and Fryer (2020) examine the relationship between charter school impacts on test scores, high school graduation, and earnings. However, they rely on selection on observables assumptions for identification so that the documented relationships may be subject to selection biases.

mathematics, English language, sciences, social studies and an essay. Students are allocated to secondary schools by the MOE based on the SEA scores and the school preferences submitted at SEA registration using the deferred acceptance mechanism summarized in Section III below.

Secondary school begins in form 1 (grade 6) and ends at form 5 (grade 10). We focus on public secondary schools of which there were 134 during our study period. Among these, there are two types of schools: Government schools (fully funded and operated by the government which enrol about 67 percent of students) and Government Assisted schools (managed by private bodies, usually a religious board, and all operating expenses funded by the government - accounting for 30 percent of enrolment).⁶ All schools provide instruction from forms 1 through 5 and teach the national curriculum. Students take two externally graded exams at the secondary level, and one at the tertiary level. The first secondary exam is the National Certificate of Secondary Education (NCSE) taken at the end of form 3 (grade 8) by all students in eight subjects.⁷ NCSE performance does not affect school progression or admission to tertiary education, and is therefore low-stakes.

The second secondary exam is the Caribbean Secondary Education Certification (CSEC) taken at the end of form 5 (grade 10) which is equivalent to the British Ordinary-levels exam. CSEC exams are given in 33 subjects. To be eligible for university admission, one must pass five or more subjects including English and mathematics. Students who qualify for university admission based on CSEC performance could either apply and, if accepted, enroll in a tertiary institution or pursue the Caribbean Advanced Proficiency Examination (CAPE). In addition, entry level positions in the public sector require at least five CSEC subject passes. For these reasons, the CSEC is a high-stakes exam. The third exam, the CAPE, is the equivalent of the British Advanced-levels exam and was launched in 2005. The CAPE program lasts two years and includes three two-unit subjects and two core subjects (Caribbean and Communication studies). Passing six CAPE units is a general admission requirement for British universities. The post-secondary qualification of a CAPE Associate's Degree is awarded after passing seven CAPE units including the two core subjects. Finally, students who obtain the highest achievable grade in eight CAPE units are awarded Government sponsored full scholarships for undergraduate studies either in Trinidad and Tobago or abroad (including the US, Canada or UK). Given this, the CAPE is a high-stakes exam.

Secondary School Applications Data: The data include the official administrative SEA covering all students who applied to a public secondary school in Trinidad and Tobago between 1995 and 2012. These data include each student's name, date of birth, gender, primary school, residential census tract, religion, SEA scores, the ranked list of secondary school choices, and the administrative school placement by the MOE. The final SEA dataset contains information on 329,481 students

⁶There were 90 Government schools and 44 Government Assisted schools during our sample period. Private secondary schools serve a very small share of the student population (about 3.4 percent).

⁷NCSE academic subjects include mathematics, English, Spanish, sciences, and social studies. NCSE non academic subjects include arts, physical education, and technical studies.

across 18 SEA cohorts. We link additional data to the SEA data by full name (first, middle, and last), gender, and date of birth.

Examination Data: To track students' exam performance and educational attainment we collected data on the NCSE exams (taken 3 years after secondary school entry, typically at age 14), the CSEC exams (taken 5 years after secondary school entry, typically at age 16) and the CAPE exams (completed after 2 years of post-secondary school studies, typically at age 18). The NCSE was launched in 2009 and data are available for years between 2009 and 2015. These data include the scores for the eight subjects assessed. The NCSE data were linked to the 2006 through 2012 SEA cohorts. The CSEC data are available for all years between 1993 and 2016. These data include the scores for each subject examination taken. The CSEC data were linked to the 1995 through 2011 SEA cohorts. The CAPE data are available for years 2005 through 2016, and are linked to the 1999 through 2009 SEA cohorts. These data contain scores for each exam unit taken.⁸

Criminal Records: We obtained the official arrests records from the Trinidad and Tobago Police Service. For each arrest that occurred in Trinidad and Tobago between January 1990 and May 2017, these data include the offender's full name, date of birth, gender, and date of arrest. To explore teen crime, these data were linked to the 1995 through 2010 SEA cohorts.

Civil Registry: We obtained the official birth records from the Trinidad and Tobago Registrar General. For each live birth in Trinidad and Tobago between January 2010 and September 2016, these data include the mother's full name, date of birth, gender, and date of the live birth. To explore teen motherhood, these data were linked to the 2004 through 2010 SEA cohorts.

Labor Market Participation: We obtained the official registry of active contributors to the national retirement fund as of May 2017 from the National Insurance Board. These data include all persons who were formally employed and, therefore, contributing to the national social security system by May 2017. For each affiliate, the data include the full current name, full original name prior to any name changes, date of birth, and gender. To explore formal employment among individuals aged 27 through 33, these data were linked to the 1995 through 2002 SEA cohorts.

Table 1 presents summary statistics for all our matched datasets. The population is roughly half female and there are about 231 students per school-cohort (column 1). About 90 percent of students took the NCSE and 73.2 percent took at least one CSEC subject. The average student passed 3.2 CSEC subjects and 34.6 percent passed five subjects including English language and math (i.e. qualified for tertiary education). We also show the outcomes by sex and the selectivity of the assigned school (by incoming SEA scores). Incoming SEA scores are 0.26 standard deviations lower for males than for females, and average scores of those assigned to the top ranked schools

⁸We matched 97.44, 96.31, and 96.6 percent of all NCSE, CSEC, and CAPE individual records to the SEA data, respectively. The non-match rate (between 3 and 4 percent) closely mimics the share of students served by private schools (3.4 percent) who would not have taken the SEA.

are 1.4 standard deviations higher than those assigned to the bottom ranked schools. Females have lower dropout rates by age 14 than males (92.1 versus 88.3 percent took the NCSE), score 0.43 standard deviations higher on the NCSE, and are more likely to qualify for tertiary education (41.5 for females versus 27.5 percent for males). Students at the most selective schools score 0.85 standard deviations higher on the NCSE than the average student at less selective schools. They also pass about 5 CSEC subjects on average, and 58.1 percent qualify for tertiary education; while this is only accomplished by 11.6 percent of students at the least selective schools (column 5).

Looking at post-secondary education, about 19.8 percent of students took at least one CAPE unit, 14.7 percent earned an Associate's degree, and only 0.95 percent earned a CAPE scholarship. Females passed 1.7 CAPE units, and 18.5 percent earned an Associate's degree. In comparison, males passed 1.1 units, and only 10.9 percent earned an Associate's degree. At the most selective schools, 33.6 percent of students took at least one CAPE unit and 25.8 percent earned an Associate's degree. Among those at less selective schools, only 4.4 percent took at least one CAPE unit and 2.3 percent earned an Associate's degree.

Moving to nonacademic outcomes, 3.3 percent of the population had been arrested by age 18. Arrests are concentrated among males of which 5.8 percent had been arrested by age 18. Arrests rates are low (1.8 percent) among students from more selective schools, and are higher (4.7 percent) among students at the least selective schools. A similar pattern is observed for teen motherhood. While 6.9 percent of girls at the top schools had a live birth before age 19, as much as 15.2 percent of females at the bottom schools did. Finally, 75.5 percent of the population is formally employed (as an adult). However, formal employment is somewhat higher for males than for females, and for those assigned to more selective schools than for those assigned to less selective schools. Next, we describe how we estimate schools' causal impacts on these key outcomes.

III Estimating School Impacts

We conceive of anything about the schooling environment that affects students as part of the school effect (or value-added) - this includes practices, facilities, teacher quality, and peers. Our first aim is to uncover the causal impact of attending each school j relative to other schools. As such, this section describes the sources of exogenous variation that we exploit for this aim, outlines the key identification assumptions, and shows empirically that these assumptions likely hold.

III.1 School Assignments

The Ministry of Education (MOE) uses a deferred acceptance mechanism to create an initial set of school assignments for students. This assignment is unrelated to unobserved student or parent actions and characteristics, so that conditional on the variables used by the assignment algorithm, the initial assignments are unrelated to unobserved determinants of student outcomes. Following

Jackson (2010), we rely on this variation to uncover schools’ causal impacts.

School assignments are as follows: Parents submit a rank-ordered list of secondary schools they wish their children to attend *before* they sit the SEA. Once the exams are scored, the top scoring student is assigned to her top choice school, then the second highest scoring student is treated similarly, and so on until all school slots are filled. Once a given school’s slots are filled, that school is then taken out of the pool, and students who had that school as their top choice will be in the applicant pool for their second choice. This process continues until all school slots are filled or all students are assigned.⁹ We refer to this rule-based initial assignment as the “tentative” assignment.

A key feature of the mechanism is that each school has a test score cutoff above which applicants are tentatively assigned to the school and below which they are not.¹⁰ Because the exact location of cutoffs is a function of the entire distribution of test scores and choices in a given cohort (which is not known to parents), the initial assignment cannot be gamed. As such, conditional on school choices and smooth functions of the SEA score, the tentative assignments are beyond parent, student or school administrator control and are therefore unrelated to unobserved determinants of student outcomes.¹¹ In reality, the official MOE placements differ from the initial assignments because principals at Government Assisted schools are allowed to admit up to twenty percent of the incoming class at their discretion.¹² This discretion is often not used by principals. However, to avoid bias, we follow Jackson (2010) and do not rely on the official MOE placement, but rather use only the exogenous variation in the tentative rule-based assignment to identify school impacts.

III.2 Identification Framework

One can write the outcome Y of student i at school j (that is, Y_{ij}) as below.

$$Y_{ij} = \nu_j + \alpha_i + v_{ij} + \varepsilon_{ij} \tag{1}$$

⁹See Appendix A for a more detailed description. Because all schools have the same preferences for students, this is similar to a serial dictatorship game. If students listed a complete ranking of all schools, this would essentially be serial dictatorship. However, because the lists are not complete, it is deferred acceptance (Gale and Shapley, 1962).

¹⁰This mechanism generates higher cutoffs for schools that tend to be higher ranked by parents so that a more preferred school will be more selective. Appendix Figure B1 shows the distribution of cutoffs for each district in the country. There is a considerable overlap of cutoffs across all districts. Indeed, Appendix Table B1 shows that all districts have schools with cutoffs below the 10th percentile and above the 75th percentile, and most have schools with cutoffs above the 90th percentile. As such, parents from all district have access to both more and less selective schools.

¹¹Note that given the realized distribution of test scores and choices, the assignment system is deterministic. However, if we consider each student’s test score and preferences to be a random draw from a distribution, then any individual pupil’s chance of being tentatively assigned to a school (which is a deterministic function of these random draws relative and their own preferences and scores) is essentially random (conditional on their test scores and choices). As such, we argue that the deterministic outcome of these random draws is exogenous to the focus family.

¹²Government Assisted schools account for about 30 percent of national enrollment. Therefore, students admitted upon discretion of principals at these schools could account at most for 6 percent of national enrollment.

In (1), v_j is the fixed “value-added” of school j to outcome Y , α_i is the fixed ability of student i , and ε_{ij} is an idiosyncratic error term. To allow for heterogeneous school impacts there is also a “match” effect between student i and school j , v_{ij} . This may be due to treatment heterogeneity along observed or unobserved dimensions, and can take any form. Average outcomes for students at school j compared to those at j' can be written as (2).

$$D_{j,j'} = [\bar{Y}_j - \bar{Y}_{j'}] = \theta_{j,j'} + A_{j,j'} + M_{j,j'} + E_{j,j'} \quad (2)$$

where, $\theta_{j,j'} \equiv [v_j - v_{j'}]$ reflects differences in value-added, $A_{j,j'} \equiv [\bar{\alpha}_{i|j} - \bar{\alpha}_{i'|j'}]$ reflects differences in the average incoming ability among individuals attending different schools, $M_{j,j'} \equiv [\bar{v}_{ij} - \bar{v}_{i'j'}]$ are differences in average match quality for the different individuals across schools, and $E_{j,j'} \equiv [\bar{\varepsilon}_{ij} - \bar{\varepsilon}_{i'j'}]$ is the difference in the idiosyncratic errors.

Application choices: Students apply to a particular ranked portfolio of schools among all possible portfolios ($c \in C$) to maximize some perceived payoff (which may be a function of match). As such, $E[M_{j,j'}|C = c] = \mu_{j,j',c}$, where $\mu_{j,j',c}$ may or may not be equal to zero.

Exogenous school assignments, conditional on choices and smooth functions of incoming test scores: For ease of exposition, we assume that all students comply with their school assignment. Students with ($C = c$) receive school assignments that are unrelated to unobserved ability conditional on smooth functions of test scores $f(SEA_i)$, so that $E[A_{j,j'}|f(SEA_i), C = c] = 0$. However, because the assignment is conditional on $C = c$, $E[M_{j,j'}|f(SEA_i), C = c] = E[M_{j,j'}|C = c] = \mu_{j,j',c}$. Intuitively, if students who choose $C = c$ have a higher match for school j than school j' , even with random assignment to schools conditional on $C = c$, there could be some differential match effects across those assigned to different schools. As such, for each $c \in C$, in expectation, the difference in outcomes conditional on smooth functions of test scores reflects true differences in value-added plus a differential match for individuals for whom $C = c$, as in (3).

$$E[D_{j,j'}|f(SEA_i), C = c] \equiv \underbrace{\theta_{j,j'}}_{\text{Difference in value-added}} + \underbrace{\mu_{j,j',c}}_{\text{Differential match for } C = c} \quad (3)$$

Due to the match term in (3), the differences in outcomes across schools for one group may not reflect differences across those same schools for those who made different choices or have different incoming test scores. However, for ease of exposition, we follow convention in this literature (e.g., Angrist et al. 2021; Mountjoy and Hickman 2020; Angrist et al. 2020), and assume constant value-added effects so that $\mu_{j,j',c} = 0$ and $E[D_{j,j'}|f(SEA_i), C = c] = \theta_{j,j'}$ for all school pairs j and j' . Aggregating across all choice groups it follows that $E[D_{j,j'}|f(SEA_i), C] \equiv \theta_{j,j'}$.

Identifying assumptions: Under the framework above, $[D_{j,j'}|f(SEA), C]$ is an unbiased estimate of the difference in fixed value-added across schools ($\theta_{j,j'}$) if: (1) school assignments are

unrelated to potential outcomes conditional on choices and smooth functions of scores, and (2) there are no differential match effects on average across schools (*this condition is satisfied under fixed value-added*). We present evidence that these two conditions hold in Sections III.6 and III.7.

III.3 Identifying Variation

Because there are multiple test-score cutoffs embedded in the assignment algorithm, the assignment mechanism generates *two* sources of exogenous variation that we can exploit (Cattaneo et al., 2020; Jackson, 2010): (a) variation around the cutoffs for each school (based on applicants scoring just above and just below each cutoff) and (b) variation across cutoffs (based on *all* students including those far away from the cutoffs). We discuss each in turn.

Variation around individual cutoffs (Discontinuity Variation):

The first source of exogenous variation is the variation around individual cutoffs. Consider the scenario illustrated in the top panel of Figure 1: Choice group 1 (left) lists school 1 as their top choice and school 3 as their second. The assignment cutoff for school 1 is 82 such that (among those in choice group 1) students who score 82 and below are assigned to school 3, while those who score above 82 are assigned to school 1. The lower left panel presents hypothetical outcome data for this group (hollow circles) who score between 70 and 90. The outcome increases smoothly with incoming test scores and there is a discontinuous jump in outcomes at the admission cutoff – coinciding to the discontinuous jump in likelihood of attending school 1 (relative to school 3) at the cutoff. With controls for the smooth relationship between incoming SEA scores and the outcome, the cutoff effect is the discontinuous jump in outcomes at the cutoff, which is the value-added of school 1 relative to that of school 3 – i.e. $\theta_{1,3}$. This is the standard variation exploited in a regression discontinuity (RD) design. This model can be implemented by regressing the outcome on smooth functions of SEA scores and an “Assigned to School 1” indicator, *using data only among choice group 1*. This variation is valid so long as the location of the cutoffs relative to the scores of students are unrelated to potential outcomes. In Appendix A we present several empirical tests to show that this condition likely holds. That is, scoring above the cutoff is not associated with a jump in density, or change in predicted outcomes but *is* strongly associated with an increased likelihood of attending ones preferred school.

Variation across cutoffs (Difference in Difference Variation):

Because there are multiple cutoffs (one for each school each year), the RD variation is not all the identifying variation embedded in the assignment mechanism. We illustrate how, under certain conditions, one can estimate the same parameter $\theta_{1,3}$ (the value-added of school 1 relative school 3) by comparing individuals away from the cutoff to those with the same test score who applied to different schools (i.e., using variation across cutoffs). To see this, consider a second set of students (choice group 2) who list school 2 as their top choice and school 3 as their second choice. School

2 has a higher cut-off than school 1 such that applicants to school 2 who score above 92 on the SEA are granted admission (top right panel of Figure 1). The lower left panel presents hypothetical outcome data for choice group 2 (black circles) who score between 70 and 90. As with choice group 1, outcomes increase smoothly with incoming test scores. However, because this group does not face a cutoff between 70 and 90, there is no corresponding jump in outcomes for this group. Students who score below 82 in both choice groups are assigned to school 3. Among these students, any difference in outcomes among students with the same test score cannot reflect a schooling effect or a test score effect, and must be attributed to differences in potential outcomes among those who make different choices. The lower left panel shows this difference as the “Choice Group Effect”. If the choice group effect and test score effects are additively separable, then the choice group effect can be estimated using the difference in outcomes across choice groups among individuals assigned to the same school (i.e., students scoring below 82) with the same incoming test score.

In the lower right panel, after accounting for the choice group effect (by subtracting it from the outcomes in choice group 2), the outcomes of both choice groups are similar among those scoring below 82 because they attend the same school, have the same incoming scores, and differences attributable to choices have been adjusted for. If choice effects and test score effects are additively separable, then the *choice-group-adjusted* outcomes for choice group 2 will approximate the counterfactual outcomes of high-scoring students in choice group 1 had they not been assigned to school 1. If so, the difference in choice-adjusted outcomes across choice groups among individuals with the same incoming scores between 83 and 92 will be roughly equal to the cutoff effect. That is, with some additional parametric assumptions, a difference-in-difference type model using variation across cutoffs can identify the same relative school effect as the RD variation within cutoffs.

By similar logic, *even though there is no cutoff between schools 1 and 2 for these two choice groups*, one can use the choice-adjusted outcomes for those in choice group 1 who score above 92 (the cutoff for school 2) to estimate the effect of attending school 2 relative to school 1 ($\theta_{2,1}$).¹³ The difference-in-difference (type) variation can be exploited by using data from multiple choice groups and then including choice group fixed effects, smooth functions of test scores, and indicator variables for being assigned to each school. This example illustrates that under additive separability of choices and test scores, and if the effect of schools is similar at all test score levels, the RD models and the DiD models will yield the same relative school estimates (see Cattaneo et al. (2020) for a general discussion of this). In Section III.6, we show that these assumptions likely hold.

¹³Intuitively, so long some students in different choice groups are assigned to the same school with the same (or similar) incoming test scores, any differences in outcomes across choice groups cannot be due to difference in school value added or incoming test scores and must therefore be due to different potential outcomes among those who make different choices. Under the additivity assumption, the choice group effect can be identified among such individuals. With the additional assumptions that the relationship between test scores is the same across choice groups and the difference in school effect is the same at all incoming test score levels, the relative value-added of all schools in the choice groups can be identified.

Making Comparisons Across all Schools:

The set of students described above allow one to estimate the relative effects for schools 1, 2, and 3 *among applicants to these schools*. However, we aim to compare the causal effect of attending each school to that of *any* other particular school. With different groups of students who make different choices (and therefore face a different set of cutoffs) one can estimate impacts for other sets of schools (say schools 3, 4, and 5). If there are considerable relative match effects and school effects are very heterogeneous by incoming achievement, then the relative effects of schools 2 and 3 for one choice group (at a given incoming achievement level) may be very different from the relative effects of those same schools 2 and 3 for a different choice group (at a given incoming achievement level). If so, it would be difficult to assess the relative impacts for schools that are not directly listed in the same choice group. However, if the relative school effects are approximately homogeneous, then school effects will be additive (i.e., the effect of attending school 1 relative to 5 is equal to the effect of attending school 1 relative to 3 plus the effect of attending school 3 relative to 5), and each school can be linked to all other schools through a chain of overlapping within-group comparisons and one can compare the effect of each school to that of every other school.¹⁴ In Section III.8 we show that this additivity condition likely holds in our setting.¹⁵

III.4 Relying Only on the Identifying Variation

We exploit variation both within and across cutoffs as discussed in Section III.3. We will refer to assigned school τ and attended school j . Based on the algorithm, students are tentatively assigned to school τ if they (a) had school τ in their list of choices, (b) scored above the cutoff for school τ , and (c) did not score above the cutoff for a more preferred school in their choices. Under the testable modelling assumptions discussed above (i.e., (1) scoring above the test score cutoffs is unrelated to potential outcomes (as shown empirically in Appendix A), (2) additive separability of test scores and school choices in determining outcomes, and (3) additivity of school effects), conditional on smooth flexible functions of incoming SEA scores and explicit controls for student choices, differences in outcomes among students with different initial tentative assignments will reflect true differences in value-added. One can therefore obtain *assigned* school's causal effects

¹⁴Suppose choice group A allows a comparison of schools 1, 2, and 3, while choice group B allows a comparison of schools 4, 5, and 6. So long as there is some other group that has one school from each group (say schools 2, 4, and 9) then all schools in 1, 2, 3, 4, 5, 6, and 9 can be compared to each-other. This example highlights that if each school can be linked to all other schools through a chain of overlapping within-group comparisons, the effects of all schools can be compared to all other schools. This identification requirement is similar to that for estimating teacher value-added while also controlling for school effects (Mansfield, 2015).

¹⁵If school effects are additive, it implies minimal match effects or treatment effect heterogeneity so that differences of effects across schools (*even those that do not have overlapping applicants*) will be equal to the relative impacts of attending one school over another for all students.

by estimating (4) by Ordinary Least Squares (OLS).

$$Y_{i\tau ct} = \Sigma(I_{i,\tau} \cdot \theta_{\tau}^{ITT}) + f(SEA_i) + \lambda_c + \mathbf{X}_{it}'\delta + S_t + \varepsilon_{i\tau ct} \quad (4)$$

In (4), $Y_{i\tau ct}$ is the outcome for student i who was assigned to school τ , and belongs to choice group c and SEA cohort t . $I_{i,\tau}$ is an indicator equal to 1 if student i was *assigned* to school τ . $f(SEA_i)$ is a smooth function of the incoming SEA score.¹⁶ X_{it} is a vector of individual-level characteristics (measured at SEA registration) including sex, district of residence fixed effects, and religion fixed effects. S_t denotes SEA cohort fixed effects; while $\varepsilon_{i\tau ct}$ is an individual-level disturbance.

Key variables for our analysis are the choice group fixed effects λ_c . These identify separate intercepts for groups of individuals who made the same school choices in the same order.¹⁷ Importantly, the choice-group indicators identify groups of individuals who may have the same SEA score but are subject to different school cutoffs – which allows for the difference-in-difference identification across cut-offs (using individuals away from the cutoff as outlined above).¹⁸ The estimated $\hat{\theta}_{\tau}^{ITT}$ s from (2) identify the average Intention-to-Treat (ITT) effect of being tentatively assigned to each school τ (relative to the same comparison school).¹⁹

III.5 Using Instruments to Obtain Causal School Attendance Impacts

To estimate Treatment-on-the Treated (TOT) effects of *attending* school j relative to other schools, we use the rule-based school assignments as instruments for actual school attendance.²⁰

¹⁶For all our main results, we model $f(SEA_i)$ with a 5th-order polynomial. However, our results are unchanged when using alternative polynomial orders (Appendix Figure B2).

¹⁷In most years, students could list four choices. However, for SEA cohorts 2001–2006 the MOE allowed students to list up to 6 different school choices (instead of the usual 4). Therefore, we grouped students with unique combinations of the first 4 choices within one set of fixed effects, and included separate sets of fixed effects for choices 5 and 6.

¹⁸In principle one could efficiently rely only on the discontinuity variation by using the “tie breaking” approach proposed in Abdulkadiroğlu et al. (2019). However, this would preclude our use of the variation away from the cutoffs which is instrumental to our ability to compare school effects across them.

¹⁹Note that because we condition on individuals rank-ordered choice lists and proximity to the school, our approach is similar to Abdulkadiroğlu et al. (2020) who assume that “any omitted variable bias afflicting OLS value-added estimates is due either to spatial heterogeneity captured by distances to each school (D_i) or to the preferences underlying the rank-ordered lists submitted to the assignment mechanism.” (Page 1513). However, unlike Abdulkadiroğlu et al. (2020) where additionally “noncompliance with the assignment mechanism, are presumed to be unrelated to potential outcomes,” we observe **both the initial assignment and the school attended**. Therefore, we **do not** rely on the additional identifying assumption of random compliance.

²⁰Noncompliance with the algorithm-based assignment may occur for two reasons. First, as explained before, principals at Government Assisted schools are allowed to replace as much as the bottom 20 percent of students tentatively assigned to their schools with any student of their choosing (see Appendix A for a detailed description of this process). The second source of noncompliance is that students may self-select and therefore not attend the schools to which they are placed. Specifically, students (and parents) may attempt to transfer to schools other than their initial placement or decide to attend a private school if they do not like their initial placement. While the first source of noncompliance is specific to the Trinidad and Tobago context, the second would exist in most contexts. As both sources of noncompliance are not random and may be related to students’ potential outcomes, this would render estimated impacts based on attended schools (without a convincing source of exogenous variation) biased.

Identification of individual school effects requires one instrument per alternative (Kirkeboen et al., 2016). We satisfy this condition in our setting by using indicators for being assigned to each school as instruments for attending each school. Ideally, all 134 schools would have strong first stages, but this is not the case. As such, to avoid being under-identified, we exclude the school assignment and attendance indicators of the schools with the weakest first stages which, therefore, serve as the omitted category in the estimation of individual school impacts.²¹ We can obtain clean causal estimates for at least one outcome for 98.5 percent of all public secondary schools in the nation (i.e., 132 schools). The resulting two-stage least squares (2SLS) model is as follows:

$$I_{i,j} = \Sigma(I_{i,\tau} \cdot \pi_{\tau j}) + f_j(SEA_i) + \lambda_{j,c} + \mathbf{X}'_{it} \delta_j + S_{j,t} + v_{ijct}, \quad \text{for each } j \in J \quad (5)$$

$$Y_{ijct} = \Sigma(\hat{I}_{i,j} \cdot \theta_j^{TOTIV}) + f(SEA_i) + \lambda_c + \mathbf{X}'_{it} \delta + S_t + \varepsilon_{ijct} \quad (6)$$

The endogenous variables are the 132 individual school attendance indicators ($I_{i,j}$) and the excluded instruments are the 132 individual school assignment indicators ($I_{i,\tau}$). We code a student as attending school j if the student was enrolled in school j at the time of writing the CSEC exams. Therefore, attended school j and assigned school τ are the same for those who comply with the exogenous assignment. While each attended school has its own assignment instrument, *all 132 school assignment indicators enter as regressors in each of the 132 first stages* denoted by (5). The $\hat{\theta}_j^{TOTIV}$ from the second stage equation (6) is an unbiased causal estimate of the effect of attending school j relative to the omitted set of schools for those who comply with the assignment. **Note:** *Because all analyses compare school estimates from the same model the particular set of schools in the omitted category does not affect any of our conclusions.*

We implement the approach outlined above to estimate individual schools' causal impacts on several outcomes. These outcomes include multiple high-stakes test scores, low-stakes test scores, school dropout, arrests by age 18, teen motherhood, and formal labor-market participation. Because we have several test outcomes, we combine similar outcomes into indexes. We created a "High-Stakes Exams" index by running a factor analysis (using the principal-component factor method) on all the CSEC and CAPE outcomes and then predicting the first unrotated factor.²² Using this same approach, we computed a "Low-Stakes Exams" index grouping both NCSE academic and non-academic performance. Appendix Table B2 shows the individual outcomes that comprise each index and the weights used to compute each index. Both indexes were standardized to have zero mean and unit variance. Other dimensions have been coded so that higher values reflect better

²¹See Appendix C for a detailed description of the school exclusion criteria.

²²The first principal component represents the maximum variance direction in the data (Jolliffe, 2002). While the sum of standardized variables has been used in other studies (e.g. Kling et al. 2007), there is no conceptual reason to put equal weight on each measure *in our context*. We prefer to use the method that best summarizes all the data. As it turns out, our measure and the sum of standardized variables has a correlation of 0.99 so that the distinction is minimal.

outcomes. These are binary indicators denoting no dropout by age 14, no live birth by age 19, no arrests by age 18, and adult (age 27+) formal labor market participation.

III.6 Testing Identification Assumption I: No Selection to Schools

As outlined in Section III.2, our two identifying assumptions are (a) no selection and (b) no differential match effects. We discuss the first here. We have already established that there is no selection in RD models using variation through admission cutoffs (Appendix A). However, as discussed in Section III.3, we do not only rely on the RD variation (at individual cutoffs) and highlight additional parametric assumptions under which we can exploit DiD variation (across cutoffs). To show that the no selection assumption also holds in our main models (which use variation both within and across cutoffs), we demonstrate that the two additional parametric assumptions under which the RD and DiD models uncover the same parameter hold, and *more importantly*, that our main estimates are consistent with what one would obtain if one used only the RD variation.

Similar School Effects by Incoming Achievement

The first parametric assumption was that the relative school effects are the same for all incoming achievement levels. Recent work (i.e., Oosterbeek et al. 2020) shows this need not hold. We test this using a re-weighting method motivated by Solon et al. (2015). Because we use variation across cutoffs, we use variation among all admitted students to compute relative school effects. Because schools admit students across a wide range of incoming scores (especially non-selective schools – see Appendix Figure B3), one can estimate relative school effects only among students with (or around) a particular incoming score. We approximate this by estimating (5) and (6) while weighting the regression on individuals at different points in the incoming test score distribution.²³ That is, where pct_i is the percentile of student i in the SEA distribution, we estimate each school’s treatment effect (θ_j^{TOTIV}) while weighting each observation by $(1 + \frac{(X - pct_i)^2}{100})^{-1}$. This puts heavy weight on students with incoming scores close to the X^{th} percentile and low weight on those far away from that percentile (see Appendix Figure B4 for the weights by SEA score). We do this for the 25th, 50th, and 75th percentiles. As pointed out in Solon et al. (2015), because weighting uncovers effects that are typical for individuals who receive more weight, the differences between unweighted and weighted estimates will be informative on the extent of heterogeneity by incoming achievement.

We document minimal heterogeneity by incoming achievement in two ways. First, we assess the extent to which unweighted and weighted estimated school effects differ using paired t -tests

²³We can only approximate this because not all schools admit students at all incoming achievement levels. See Appendix Figure B3 for the distributions of incoming test scores by school selectivity. While almost all schools admit students around the 75th percentile, effects weighted at that percentile will be largely representative of effects for students close to that percentile. However, because the most selective schools do not admit students at the 25th percentile, school effects weighted at this percentile will be largely representative of effects for students close to that percentile for the least selective half of schools, but will reflect marginal effects for the lowest-scoring admits for schools that admit student exclusively above this achievement level.

(Appendix Table B3).²⁴ Based on 1,399 pairwise t -tests (by school-outcome-percentile), in less than 2 percent of cases does one reject that the estimates are the same (at the 5 percent significance level).²⁵ Second, we show that the unweighted estimates are very good predictors of the weighted ones. In Appendix Figure B5, we plot the weighted effects (and the prediction 95% confidence interval) against the unweighted (or average) school impacts. While there is some evidence that the average effects are not always the same as those evaluated at the 75th percentile, they are very similar so that for all three percentiles, the average-weighted effects are close to the 45° line.²⁶ As such, our assumption that school effects are the same by incoming achievement is largely supported by the data. However, because school effects may not be identical throughout the entire incoming score distribution, we show that our conclusions still hold when (a) limiting the analysis to a homogeneous set of students, and (b) conducting the analysis among subgroups of students among which there is strong evidence of homogeneous effects. **Note**, moreover, that we will show that our main estimates are consistent with those that use the RD variation only.

Robustness to Interactions

The second assumption was that there are no interaction effects between incoming SEA scores and school choices. A key difference between a model that uses the DiD variation (as we do) and one that relies only on variation at cutoffs (not across) is that our model excludes interactions between school choices and incoming test scores. If the interactions are important for identification, then our school estimates (using the DiD variation without interactions) will be biased relative to flexible models that rely only on variation at cutoffs by accounting for interactions. As such, one can test the importance of the additive separability assumption between school choices and incoming scores by seeing if our estimates are robust to the inclusion of interactions between test scores and school choices. Note that we cannot fully control for all the possible interactions (because this is the level of the variation). However, the stability of our estimates to the inclusion of interactions between coarse measures of test scores and choices may be informative. We, therefore, estimate models controlling for these interactions and compare the resulting estimated school impacts, $\hat{\theta}_j^{TOTIV}$, to those that we obtained without these extra controls (reported in Appendix D). In all cases, the correlations between the resulting effects is close to 1. That is, consistent with the test

²⁴Note that we implement a conservative estimate of the standard error of the differences by assuming that the correlation between the errors for the two estimates of the same school is 1.

²⁵Looking at individual outcomes, for all outcomes one cannot reject equality of effects between the average and those evaluated at the 25th percentile or the median *for any school*. The only estimates that appear to differ are for dropout and arrests evaluated at the 75th percentile. Even among these, the vast majority are the same – but insofar as out average effect may not reflect estimates for all students, it would only be for a small number of school effects for only these two outcomes among the highest achieving students.

²⁶The slope predicting the percentile-weighted effects using the average effect is 0.998 for the 25th percentile, 0.923 for the 50th percentile, and 0.707 for the 75th percentile. Only for the 75th percentile does one reject that the slope is equal to one at the 5 percent level.

above, our estimates are likely the same as those using the RD variation alone.

Regression Discontinuity Variation vs. All Variation

While the tests above indicate that the assumptions required for the RD and DiD models to yield the same effects are largely satisfied, they do not directly show that our estimates are the same as an RD model - which is the condition we require. Here, we validate our school estimates (that use variation both within and across cutoffs) using only local RD variation through individual cutoffs *that do not exploit any variation across cutoffs*. If the results using all the variation are similar to those obtained using only the RD variation right around individual cutoffs, it would indicate that either (a) the additional DiD identifying assumptions are satisfied, or (b) any violations of the DiD identifying assumptions are not severe enough to lead to appreciable bias. Indeed, we show this.

For each school in each SEA cohort, we estimate the RD effect of scoring above the rule-based assignment cutoff for that school in that cohort. As pointed out in Kirkeboen et al. (2016), this is the difference in school quality between attending the preferred school versus attending the set of counterfactual schools for those applying to that school in that cohort. We also obtain an estimate of the effect of scoring above the cutoff on the impact of the attended school (i.e., $\hat{\theta}_j^{TOTIV}$) relative to that of *the same set of counterfactual schools*.²⁷ The actual cutoff effect is the change in actual outcomes evaluated at a particular cutoff with flexible functions of test scores specific to that cutoff (i.e., not relying on variation across cutoffs). As such, if our school IV estimates reflect the causal impact of attending school j relative to the counterfactual schools, then the RD effect on actual outcomes should be similar to the RD effect on our estimated IV school impacts.²⁸

To show this, we regress the RD effects on actual outcomes on the RD effects on the estimated IV school impacts and test if the slope is statistically distinguishable from 1. We first implement this test using the raw school effects. Following Hastings et al. (2015), to account for noisiness in the RD estimated effects, we weight each RD estimate by the inverse of its squared standard error. Additionally, to account for estimation errors in the RD effects on school impacts, we implement

²⁷All RD estimated effects of scoring above the cutoff for each school in each year use the optimal bandwidths derived from Imbens and Kalyanaraman (2012).

²⁸A similar test was implemented in Hastings et al. (2015) and Beuermann and Jackson (2020). This is also similar in spirit to the random assignment validation of school value-added in Deming et al. (2014). See Appendix D for further discussion of this test.

Empirical Bayes estimates of each cutoff effect.²⁹ The binned scatter-plot, pooled across outcomes (Figure 2) presents the relationship for the raw school impacts (left) and the Empirical Bayes estimates (right) – the conclusions are the same. Using the raw estimates, we fail to reject that the estimated slope coefficient is different from 1 (p -value = 0.48). The slopes using the Empirical Bayes estimates (with or without weights) are close to 1, and both yield p -values well above 0.5. This test (in addition to the two previous tests) indicate that the "no selection" condition which we showed holds for the RD variation likely also holds for our preferred models.

III.7 Testing Identification Assumption II: No Differential Match Effects

Now we focus on the second identifying assumption. While we assumed that value-added is fixed for exposition purposes, here we test the specific identifying assumption we require: no differential match effects on average across schools. Following Kirkeboen et al. (2016) we test for differential match bias within choice groups (what they refer to as comparative advantage) and assess whether students who list school m over school k experience larger benefits from attending school m versus school k than those who ranked school k over school m . To this aim, for every pair of schools m and k , we estimate (7) among individuals who list both schools in their choices. All common variables are defined as before.

$$Y = I_m \cdot \theta_{m,k}^{ITT} + (I_m \cdot I_{k-b-m}) \cdot \theta_{k,m}^{Pref} + I_{\neq m,k} \cdot \theta_{n,k}^{ITT} + f(SEA) + \lambda_c + \mathbf{X}'\delta + S_t + \varepsilon \quad (7)$$

In (7), I_m is an indicator denoting being assigned to school m , and $I_{\neq m,k}$ is an indicator denoting whether the individual was *neither assigned* to school m nor school k . Because the omitted school assignment is school k , $\theta_{m,k}^{ITT}$ (the coefficient on I_m) captures the effect of being assigned to school m relative to school k . The variable I_{k-b-m} is an indicator that denotes individuals who list school k above m (as opposed to school m above k). The direct effect of this difference in preferences is already accounted for in λ_c . The interaction effect between being assigned to school m relative to k and listing school k above m (i.e., $\theta_{k,m}^{Pref}$) captures the *differential* effect of being assigned to school m relative to k for those who chose k before m relative to those who listed m before k .

²⁹Specifically, for any particular outcome, the predicted RD effect for school j is the weighted difference in estimated school impacts between those just above and below the cutoff for school j . We can express the estimated parameter as $\hat{\zeta}_{\hat{\theta}_j} = \sum_{k \in A} a_k(\hat{\theta}_k) - \sum_{k \in B} b_k(\hat{\theta}_k) = \sum_{k \in A} a_k(\theta_k + \varepsilon_k) - \sum_{k \in B} b_k(\theta_k + \varepsilon_k)$, where A is the set of schools the students attend above the cutoff, B the set of schools the students attend below the cutoff, a_k and b_k the proportions in which they do so and ε_k the estimation error for the school impact θ_k . If we assume that the school impacts are not independent within each cutoff, but that the estimation errors are, we can approximate the variance of this estimated parameter as $Var(\hat{\zeta}_{\hat{\theta}_j}) = \sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta} + \sum w_k SE_k^2$, where σ_{θ}^2 is given by the magnitude of the school impacts, Cov_{θ} is approximated by the covariance between each pair of schools the students applied to and SE_k^2 is given by the square of the standard error of the school impact. If we also include the squared standard error of the RD estimate, $SE_{\hat{\zeta}}$, the reliability ratio of our RD estimate is given by $\lambda_j = \frac{\sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta}}{(\sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta}) + (\sum w_k SE_k^2 + SE_{\hat{\zeta}}^2)}$. Therefore, our Empirical Bayes estimate of the *predicted* effect of cutoff j is $[\lambda_j \times \hat{\zeta}_{\hat{\theta}_j}]$.

If, for any pair of schools, students who prefer school m over k experience larger relative effects from attending school m over k , then $\theta_{k,m}^{Pref}$ would be negative. We estimate this model for all pairs of schools, and capture $\hat{\theta}_{k,m}^{Pref}$ for each pair. We then test for whether the average and the median of these estimates differ from zero.³⁰ We find small average and median effects failing to reject that these are equal to zero at the 10 percent level (see [Appendix Table B4](#)) – suggesting minimal differential match effects of this sort.

III.8 Testing for Additivity of School Effects

Finally, as detailed in [Section III.3](#), even if we can identify relative school effects among individuals who make similar choices (as shown in [Sections III.6](#) and [III.7](#)), our ability to compare *all* schools against each other relies on the assumption that school effects are additive. This condition holds with fixed value-added, but we show that this particular condition holds empirically. We define the parameter $\theta_{m,k}^{ITT}$ as the effect of being assigned to school m relative to being assigned to school k . If school effects are additive, then $\theta_{m,k}^{ITT} = \theta_{m,l}^{ITT} + \theta_{l,k}^{ITT}$. If school effects are not additive, then this condition will generally not hold. As such, to test for additivity, we implement the sample analog of this test. That is, we test whether $\hat{\theta}_{m,k}^{ITT} = \hat{\theta}_{m,l}^{ITT} + \hat{\theta}_{l,k}^{ITT}$. Specifically, for each pair of schools m and k , we restrict the data to students that had both schools in their choices. We then estimate equation (8), where I_m is an indicator for being assigned to school m , and $I_{\neq m,k}$ is an indicator connoting assignment to a school other than school m or school k .³¹

$$Y = I_m \cdot \theta_{m,k}^{ITT} + I_{\neq m,k} \cdot \theta_{n,k}^{ITT} + f(SEA) + \lambda_c + \mathbf{X}'\delta + S_t + \varepsilon \quad (8)$$

In equation (8), because the omitted school assignment is school k , $\theta_{m,k}^{ITT}$ captures the effect of being assigned to school m relative to school k . We then find all intermediate schools l such that $\hat{\theta}_{m,k}^{ITTsum} = \hat{\theta}_{m,l}^{ITT} + \hat{\theta}_{l,k}^{ITT}$ can be computed. As is typical in the value-added literature, because we will use this sum of estimates as a regressor, we form Empirical Bayes estimates by multiplying each raw sum of effects $\hat{\theta}_{m,k}^{ITTsum}$ by an estimate of their reliability $\hat{\lambda}_{m,k}$.³² That is, our Empirical

³⁰Because we do not account for estimation errors, this test is biasing *toward* rejecting the null of no match effects.

³¹Estimated standard errors are clustered at the assigned school level.

³²Where $\sigma_{\theta_{sum}}^2$ is the variance of the sum of the two school effects to create the indirect estimate, and $\sigma_{e,sum}^2$ is the variance of the estimation error for the indirect effect, the reliability ratio of our indirect school effect can be written as: $\lambda_{mk} = \frac{\sigma_{\theta_{sum}}^2}{\sigma_{\theta_{sum}}^2 + \sigma_{e,sum}^2}$. To compute an estimate of the reliability ratio, we assume that the estimation errors are uncorrelated across schools. Therefore, our empirical estimate of $\sigma_{e,sum}^2$ is $(SE_{\theta_{ml}}^2 + SE_{\theta_{lk}}^2)$. We then make the additional standard assumption that the estimation errors are uncorrelated with the true effects. Therefore, the total variance of the estimated indirect school effects is $\sigma_{\hat{\theta}_{sum}}^2 = \sigma_{\theta_{sum}}^2 + \mathbf{E}(SE_{\hat{\theta}})^2$, where $\mathbf{E}(SE_{\hat{\theta}})^2$ is the sum of the squared expected standard errors $SE_{\hat{\theta}_{ml}}$ and $SE_{\hat{\theta}_{lk}}$, which are approximated by the mean standard error of the estimates (alternatively, we also approximated them with the median standard error of the estimates yielding equivalent results). By subtracting our estimates of $\mathbf{E}(SE_{\hat{\theta}})^2$ from the overall variance of the estimated indirect effects, we estimate $\sigma_{\theta_{sum}}^2$. The estimated reliability ratio, $\hat{\lambda}_{m,k}$, is then computed using these two estimates ([Kane and Staiger, 2008](#)).

Bayes estimate of the indirect effect is $\hat{\lambda}_{mk} \hat{\theta}_{m,k}^{ITT_{sum}}$. Using (9) we test whether $\beta = 1$, to implement a formal test of additivity of school effects.

$$\hat{\theta}_{m,k}^{ITT} = \alpha + \beta [\hat{\lambda}_{mk} \hat{\theta}_{m,k}^{ITT_{sum}}] + \varepsilon \quad (9)$$

We implement this test on sets of school pairs that have at least 100 observations assigned to school m , school k , or school l across all SEA cohorts – this corresponds to about 15 observations per school per cohort. A binned scatterplot of the direct estimates against the indirect estimate (i.e., the sum) is presented in Figure 3. The datapoints line up remarkably well along the 45 degree line, and the null hypothesis that the effects are additive (i.e. $\beta = 1$) is not rejected. This suggests that our school effects are additive – indicating that the fixed value-added assumption is reasonable.

IV Magnitude of the School Impacts

To assess the magnitude of the school impacts on each outcome, we estimate the standard deviation of these impacts. Because the school effects are estimated with error and noise, simply reporting the variance of the estimated effects would overstate the magnitude of schools’ actual impacts. As is common practice in the school and teacher effects literatures (e.g., Kane and Staiger 2008; Chetty et al. 2014), to account for this, we rely on the correlations between school effects across years to identify the variance of persistent school effects. Following Jackson (2013), we do this in two steps. First we estimate the IV impacts of each school with two different sub-samples. One comprising even SEA cohorts ($\hat{\theta}_{j,even}^{TOT_{IV}}$) and other including odd cohorts ($\hat{\theta}_{j,odd}^{TOT_{IV}}$). Let $p \in \{\text{even, odd}\}$. These two school estimates contain a persistent school effect (θ_j^{TOT}) and a transitory effect (μ_{jp}). In a second step, under the assumption of joint normality of these components and the covariance structure in (10), we uncover Maximum Likelihood estimates of the variance of the persistent school impacts ($\sigma_{\theta_{TOT}}^2$) and of the transitory school impacts (σ_{μ}^2).³³

$$\begin{bmatrix} \theta_j^{TOT} \\ \mu_{jp} \end{bmatrix} \sim N \left(0, \begin{pmatrix} \sigma_{\theta_{TOT}}^2 I_J & 0 \\ 0 & \sigma_{\mu}^2 I_{(J \times p)} \end{pmatrix} \right) \quad (10)$$

Table 2 reports estimates of the standard deviation of the persistent school impacts for each outcome along with their 90 percent confidence intervals estimated by bootstrap (column 1).³⁴ To aid interpretation, all outcomes are coded so that higher values reflect better outcomes.

High-stakes exams: The persistent school effects for the high-stakes dimension have a standard

³³The key assumption here is that the error terms are uncorrelated across even and odd cohorts. As such, the covariance between the even and odd effects reflects the variance of the persistent effect.

³⁴To be conservative, we exclude outlier schools with estimates lying 4σ away from the median. In Appendix Table B5 we also show estimates of the standard deviation of the persistent school impacts without removing outliers; as well as estimates using weighted school effects. All estimates are qualitatively similar.

deviation of 0.441 (with a 90% confidence interval between 0.4 and 0.48). This indicates that attending a school at the 85th percentile of the impact distribution compared to attending a school at the median (for five to seven years) would increase high-stakes test performance by approximately 0.44 standard deviations. These estimated school impact sizes are larger than those found for school impacts on test scores in North Carolina (Jackson 2013; Deming 2014), and than those of attending Promise Academy in the Harlem Children’s Zone (Dobbie and Fryer, 2015); but on the same order of magnitude as that of attending high-impact Boston urban charter schools (Angrist et al., 2013).

Low-stakes exams: The magnitude of the school impacts on high-stakes and low-stakes tests are very similar. The standard deviation of the persistent school effect on the low-stakes index is 0.473 (with a 90% confidence interval between 0.43 and 0.51). That is, attending a school at the 85th percentile of the impact distribution compared to a school at the median (for three years) would increase low-stakes test performance by approximately 0.47 standard deviations.

Dropout: Because all students take the NCSE exams around age 14, our measure of dropout is not being registered for the NCSE exams. The estimated standard deviation of the persistent school impacts is 0.09 – indicating that attending a school at the 85th percentile of the impact distribution compared to attending a school at the median would reduce high school dropout by approximately 9 percentage points. Our estimated impact of attending a school with 1σ higher impact on dropout is similar than that of attending a charter high school (Booker et al., 2011) or winning a lottery to a choice school in North Carolina (Deming et al., 2014). As such, our estimates are in line with what one might expect based on existing studies.³⁵

Crime: The standard deviation of the persistent school effects is 0.037, which means that being assigned to a school at 85th percentile of the impact distribution as opposed to the median would reduce the likelihood of being arrested as a teenager by 3.7 percentage points. Relative to the average arrest rate of 3.3 percent, this is a sizable reduction in teen arrests. Our estimates are larger than, but generally in line with Deming (2011) who finds that winning a lottery to attend a better school reduced arrests among high-risk youth by about fifty percent. The confidence interval does not include zero so that these school impacts are real and persistent over time.

Teen motherhood: The standard deviation of the persistent school effects on teen motherhood is 0.173. The 90 percent confidence interval is between 0.15 and 0.19. Going from a school at the median to one at the 85th percentile of the impact distribution would reduce teen live births by 17.3 percentage points. While there are many studies of the impact of teen motherhood on schooling, we believe that this is the first study to examine the distribution of individual schools’ causal impacts

³⁵The charter school and choice school literatures find impacts on high school completion between 10 and 15 percentage points. Our estimates suggest that these choice schools may be more than 1σ above the typical school.

on teen motherhood.³⁶ Given that the teen live birth rate is around 10 percent on average, these represent large economically important relative impacts.

Labor market participation: The final outcome we examine is participating in the formal labor market. We examine school effects on the likelihood that a student is observed with positive earnings in the formal labor market (i.e. contributing to the national social security system). The standard deviation of the persistent school effects on this outcome is 0.07 (with a 90% confidence interval between 0.06 and 0.08). Going from a school at the median of the impact distribution to one at the 85th percentile would increase the likelihood of being formally employed by 7 percentage points. This impact is economically meaningful.

The fact that schools have economically meaningful impacts on an array of different outcomes is not surprising. However, the policy implications of this result depends on the extent to which these school impacts are all well-measured by a school’s impact on high-stakes exams. If school impacts across these outcomes are highly correlated, then school impacts on high stakes exams would identify those schools that will improve life outcomes. Using these estimates to inform policy (such as allocating funds, school closures, or rewards) would likely improve all outcomes. However, if those schools that improve high-stakes exams are a different set of schools than those that improve labor market participation or those that reduce crime, it would mean that commonly used test-based measures of school quality are incomplete. In such a scenario, using school impacts on high stakes exams to inform policy could have deleterious impacts on other outcomes and could lead to multitasking problems (Holmstrom and Milgrom, 1991). We examine the relationship between school impacts across these different outcomes below.

IV.1 Is School Quality Unidimensional?

Many recent education policies (e.g., **No Child Left Behind** in the U.S. or **League Tables** in the U.K.) are predicated on the idea that schools that raise test scores are better schools. While this may be true *on average*, if school quality is multidimensional, school impacts on test scores may not capture impacts on other important dimensions of quality. To assess this, we explore the relationship between estimated school impacts on high stakes test scores and other outcomes. To avoid attenuated correlations due to estimation errors in the estimated school effects, we obtain Maximum Likelihood estimates of the true correlations between each pair of outcomes (as in Abdulkadiroğlu et al. 2020).³⁷ Following the notation in equation (10), consider two outcomes 1 and 2 so that θ_{1j}^{TOT} is the persistent effect of school j on outcome 1 and θ_{2j}^{TOT} is the persistent effect

³⁶In related work, Jackson (2019) finds that converting existing coeducation school to single-sex reduced the teen birth rate by 4 percentage points. Dobbie and Fryer (2015) find that females admitted the a charter school in Harlem Children’s zone are 10.1 percentage points less likely to be pregnant in their teens. Beuermann and Jackson (2020) find that attending a preferred school in Barbados decreases the teen motherhood rate by about 6 percentage points.

³⁷The scatterplots of the raw school effects are shown in Appendix Figure B6

of school j on outcome 2. Similarly, μ_{1jp} and μ_{2jp} are the transitory effects of school j in period p on outcomes 1 and 2. Under the assumption that the effects on outcomes 1 and 2 follow a joint normal distribution as in equation (11), one can estimate the correlation (net of estimation errors) between the effects on any two outcomes 1 and 2, (that is, ρ_{12}), by Maximum Likelihood.³⁸

$$\begin{bmatrix} \theta_{1j}^{TOT} \\ \theta_{2j}^{TOT} \\ \mu_{1jp} \\ \mu_{2jp} \end{bmatrix} \sim N \left(0, \begin{pmatrix} \sigma_{\theta_1}^2 I_J & \rho_{12} \sigma_{\theta_1} \sigma_{\theta_2} I_J & 0 & 0 \\ \rho_{12} \sigma_{\theta_1} \sigma_{\theta_2} I_J & \sigma_{\theta_2}^2 I_J & 0 & 0 \\ 0 & 0 & \sigma_{\mu_1}^2 I_{(J \times p)} & 0 \\ 0 & 0 & 0 & \sigma_{\mu_2}^2 I_{(J \times p)} \end{pmatrix} \right) \quad (11)$$

Table 2 presents correlations between equally weighted (or average) school impacts on high stakes tests and the other outcomes (column 2). School impacts on high-stakes tests do not explain large shares of school effects on the other outcomes. The correlation between school impacts on the high-stakes and low-stakes exam indexes is only 0.1 (with a 90% CI that includes zero). While this may seem low, recall that in addition to the difference in stakes, the low stakes exams include non-academic subjects such as physical education and arts. Indeed, the correlation between school impacts on the high-stakes exams and the low-stakes *academic* exams is 0.2, while that for the low-stakes non-academic exams is 0.01 (Appendix Table B5). The correlation with dropout is 0.12, and with being formally employed is 0.15. This suggests that schools that improve high-stakes exams are associated with relatively small improvements in these other outcomes. The correlations between performance on high-stakes and arrests are moderate and positive (0.28), while that with no teen motherhood is slightly negative and statistically indistinguishable from zero. This suggests that schools that improve high-stakes exams performance also tend to reduce arrests, on average, but that only about 7.8 percent (i.e. $0.28 \times 0.28 = 0.078$) of the variation in school impacts on reduced arrests can be explained by effects on high-stakes exams, and *vice versa*. While this may seem low, a disconnect between school impacts on high-stakes and other outcomes like low-stakes exams and crime has been documented in other settings (e.g. Mbiti et al. 2019; Deming 2011).

Correlations of School Effects Across Outcomes by Incoming Achievement:

One may worry that the low correlations for school impacts across outcomes *may* reflect students who are marginal for different outcomes attending different schools. For example, suppose only low-achieving students are marginal for high school dropout, while only high-achieving students are marginal for high-stakes exams. Also, suppose that school A only admits high achievers

³⁸Intuitively, the raw correlation can be uncovered based on the correlation between even SEA cohorts for one outcome and odd cohorts for another, and *vice versa*. To dissattenuate this raw correlation one must divide by the square root of the product of the reliability ratios for each measure (Spearman, 1904). The reliability of each measure is obtained using the correlation between even and odd year estimates for that same outcome. Doing this calculation manually yields very similar results (see Appendix Table B6).

and B low achievers. Even if both schools have the same *potential* impacts on dropout and high-stakes exams, school A will appear to improve high-stakes exams while school B will only appear to influence dropout.³⁹ To assess this possibility, we use the weighted school impacts (as discussed in Section III.6) to obtain the Maximum Likelihood estimates of the correlations of impacts across outcomes for students at different points in the achievement distribution. By weighting the school impacts at the same point in the achievement distribution *for all outcomes*, we ensure that the resulting correlation of impacts across outcomes is not due to differences in the students being compared, but rather due to differences in school effectiveness (for the same set of students). [Appendix Figure B3](#) shows that roughly 90 percent of all schools admit students close to the 75th percentile, so that there is common support in this range. As such, school effects weighted around the 75th percentile are comparable across schools and represent students that are quite similar – allowing us to estimate correlations of school effects across outcomes for students that are largely similar.

We present the Maximum Likelihood estimated correlations between the estimated school impacts on high-stakes and other outcomes when we center the weights at the 75th percentiles of the SEA distribution in [Table 2](#) (column 3).⁴⁰ Generally, even among students with the same incoming level of achievement, the correlations between schools high stakes exams impacts and those on other outcomes are low. Using weights centered on the 75th percentile increases the correlations with teen arrests to 0.41, but all other correlations remain below 0.2. Taken together, the results do not support the notion that these low correlations are mainly due to different schools serving students who are marginal for different outcomes – suggesting that the low correlations reflect different schools having impacts on different outcomes.

The patterns in [Table 2](#) indicate that (a) schools have economically meaningful impacts on a range of outcomes and (b) impacts on these different dimensions are not strongly related. This suggests that school impacts on no single outcome serves as a “summary measure” for school quality. As such, the extent to which parents choose different schools for their children may have to do with the extent to which they value school impacts on different dimensions. We showed that school impacts on non-academic dimensions are economically meaningful and large. As such, the fact that parents may not choose schools that improve test scores (e.g. [Abdulkadiroğlu et al. 2020](#); [MacLeod and Urquiola 2015](#)) may reflect parents choosing schools that improve other outcomes

³⁹Imagine two schools. Both schools have effect 1 on high stakes and 1 on dropout. However, only students below 60 have a dropout effect and only student above 40 have a high stakes effect. School 1 ability is uniformly distributed between 0 and 60. In this scenario, the effect on high-stakes for the average admit to school 1 is [1 for 33% of students above 40] = 0.33. Also, the effect on dropout for the average admit to school 1 is [1 for all students] = 1. In contrast, School 2 ability is uniformly distributed between 40 and 100. As such, the effect on high-stakes for the average admit to school 1 is [1 for all 100% of students] = 1. Also, the effect on dropout for the average admit to school 1 is [1 for 33% of students below 60] = 0.33. In this highly stylized example, the correlations between the effects would be -1 even though they would have the same effect had they admitted the same students.

⁴⁰In [Appendix Table B5](#) we also show estimated correlations when we center the weights at the 25th and 50th percentiles of the SEA distribution. Similarly, all estimates suggest moderate to low correlations.

(that are weakly related to test score impacts). We explore these possibilities in Section V.

V Estimating Preferences for Schools

In this section we will use the estimated school impacts and the set of secondary school choices to (a) examine the extent to which parents choose schools based on their causal impacts, and (b) explore the extent to which they choose schools with casual impacts on outcomes other than high-stakes tests. As in all studies of this type (e.g. Avery et al. 2013; Burgess et al. 2015; Abdulkadiroğlu et al. 2020; Hastings et al. 2006), we infer that parents “value” or “prefer” schools that they rank more highly. However, *we cannot observe preferences for school attributes directly*. Rather, by observing the attributes of preferred schools we can assess the extent to which preferences for schools are correlated with particular school attributes. As such, even though the relationships between school preferences and school attributes we present are robust across several models and to the inclusion of a rich set of controls, as in other studies of parent choices we cannot entirely rule out that unobserved determinants of parent choices may affect our results.

Most discrete choice models infer preferences under the assumptions that the top ranked choice is the most preferred option of all options, the second is the second preferred, and so on. When choices are unlimited and assignments are based on deferred acceptance it is rational for individuals to make choices in this way. Accordingly, assuming rational choices, preferences can be inferred using the standard discrete choice models. However, under deferred acceptance assignment with a limited set of choices (as in our setting and in many others), these choices may be strategic such that the top-ranked school *may not be* the most preferred option of all options, and the second *may not be* the second preferred, and so on. Because more desirable schools will tend to have higher admission cutoffs (Jackson 2010), this kind of strategic choice is quite likely to occur among low-scoring applicants who can only feasibly attend a smaller number (and less desirable set) of schools than higher-scoring applicants. As such, standard models that assume the choices are truthful to infer preferences may not be appropriate for our setting. Instead, we propose a modification to the standard multinomial logit model that does not assume, *or require*, truthfully revealing choices. Similar to Agarwal and Somaini (2018) we do not take the rank-order lists as true preferences, but rather assume that the submitted list is an optimal choice of a lottery over possible school assignments. Our model accounts for rational strategic behaviours explicitly, allowing one to infer preferences for schools so long as choice behaviours are rational.

V.1 A Model of School Choices

To model school choices, we make a distinction between the *ex post* utility of attending a school and the *ex ante* utility of *applying to* a school. We rely on theoretical results about rational choices under deferred acceptance vis a vis these two concepts to infer preferences for schools. For ease

of exposition, we assume one parent per child. We derive the choice probability from the utility-maximizing behavior of parent $i \in N$, on behalf of student $i \in N$. Parents choose a finite number (R) of schools among all schools in the nation. Each school is indexed by $j \in J$. The ex-post utility parent i derives from student i attending each school alternative j has the following general form:

$$U_{ij} = U(X_i, Z_{ij}, \varepsilon_{ij}) = \delta(X_i, Z_{ij}) + \varepsilon_{ij} \quad (12)$$

where $U(\cdot)$ is the function mapping school attributes and student characteristics to utility values U_{ij} , X_i are observed student characteristics, Z_{ij} are observed school-specific attributes that may vary at the student i level (such as proximity to primary school), and ε_{ij} is a random error.

The school choice set is the same for all parents (i.e., $J_i = J \forall i$), and each parent submits a single ranked-ordered list. Let $U_{ij}^{r_i s}$ indicate the utility parent i gets from school j that they ranked in position s ($r_i = s$), so that $U_{ij}^{r_i 1}$ is their utility for the school ranked first, $U_{ij}^{r_i 2}$ is their utility for the school ranked second, and so on. Let $U_{ij}^{r_i 0}$ indicate the utility parent i gets from attending school j that they did not rank. Under the algorithm used to assign students to schools, among the ranked schools, parents have incentives to truthfully reveal their preference rankings (Haeringer and Klijn 2009; Pathak and Sönmez 2013). If parents make rational choices then condition (1) below holds:

Condition (1): $U_{ij}^{r_i a} > U_{ik}^{r_i b} \forall k \neq j \in J, a < b$ and $b \neq \emptyset$: Parent i prefers their a -ranked school over any other school k ranked below.

One could rely only on comparisons within the set of submitted choices to infer preferences about schools (e.g. Avery et al. 2013; Beuermann and Jackson 2020). However, if not all schools are ranked, comparisons made only among chosen schools can *potentially* be misleading about particular attributes if the set of choices is not random. To see this, imagine that all parents chose four schools that are very close to home. If one were to look only within the set of schools listed, one might infer that proximity is unrelated to choices when the opposite is true. To avoid this problem, one must compare choices made (or at least one of the choices made) against *all* possible choices (e.g. Hastings et al. 2009), or make assumptions about the set of options that could have been chosen (e.g. Abdulkadiroğlu et al. 2020). We follow the less restrictive approach and compare the top choice to *all* the un-chosen schools (while explicitly accounting for proximity).

When parents are unconstrained in the number of schools they can list, then the top choice is the most preferred school of all possible schools, that is $U_{ij}^{r_i 1} > U_{ik} \forall k \neq j \in J$ (Roth and Oliveira Sotomayor, 1992).⁴¹ However, when the number of choices is limited, parents may act strategically so that the top listed choice is not necessarily the school they *ex-post* would prefer. Chade and Smith (2006) demonstrate that when the number of choices is limited, it is rational for parents to

⁴¹This condition is often assumed without testing it explicitly even when choices are constrained.

maximize the expected value of the set of choices, where the expected value of applying to a set of schools is a function of both the ex-post utility of attending the listed schools and the likelihoods of being admitted to those schools.⁴² When listing a finite set of schools, it is rational to trade-off the ex-post utility associated with attending a school against the probability of being admitted. As shown in [Chade and Smith \(2006\)](#), if a parent's ex-post most preferred school (i.e. the school with the highest U_{ij}) is not the top choice, it must be because the probability of admission to that ex-post preferred school is too low. As such, the top choice school may not be the school with the highest ex-post utility of attendance, but will be the school the highest ex-post utility *given the probability of admission*. A useful empirical prediction from [Chade and Smith \(2006\)](#) is that *with strategic choices* so long as the parents are rational, conditional on the probability of admission, the top choice school must have higher ex-post utility than any unranked school. Where p_{ij} is the probability that student i is admitted to school j , this yields condition (2) below:

Condition (2): $U_{ij}^{r_{i1}} | p_{ij} > U_{ik}^{r_{i0}} | p_{ik} \forall k \neq j \in J$: Conditional on the admission probabilities, parent i ex-post prefers their first-ranked school over any school not in the submitted choices.

The two conditions suggest that, where R_i is the maximum number of alternatives ranked by parent i , assuming rational choices, the probability that a parent i submits a particular ranking over all schools is given by equation (13) below.

$$\begin{aligned}
 Prob[r_{i1}, r_{i2}, \dots, R_i] = Pr[& \underbrace{(U_{ij}^{r_{i1}} | p_{ij} > U_{ik}^{r_{i0}} | p_{ik} \forall k \neq j \in J)}_{\text{Top choice preferred to all non-chosen schools conditional on admission probabilities}} \\
 & \cap \underbrace{(U_{ij}^{r_{i1}} > U_{ik}^{r_{im}}, 1 < m, \forall m \in \{2, \dots, R_i\}) \cap \dots \cap (U_{ij}^{r_{iR_i-1}} > U_{ik}^{r_{iR_i}})}_{\text{Higher ranked chosen ex-post schools are preferred to lower-ranked chosen schools}}] \quad (13)
 \end{aligned}$$

Therefore, if one had measures of the admission probabilities, and one correctly modelled how admission probabilities influence choices, then one can infer *ex-post* preferences across all schools based on the choices (even when the top choice is not the ex-post preferred school).

V.2 Modified Exploded Multinomial Logistic Model

Equation (13) defines the likelihood of observing a set of choices as a function of parent utilities for schools and random errors. We make some assumptions on the form of U_{ij} , the form of $U_{ij} | p_{ij}$, and the distribution of ε_{ij} to model equation (13) and use the observed choices to infer parental preferences for school attributes. Following [Hastings et al. \(2009\)](#) and [Abdulkadiroğlu et al. \(2020\)](#), we parametrize $\delta(\cdot)$ as a linear-in-parameters function of the school characteristics. Where β is

⁴²When there are finite choices, as in our setting, rational agents will choose the portfolio of four schools that as a whole provide the greatest expected utility. Once this set of schools is decided, they will order them by ex-post utility (see [Chade and Smith \(2006\)](#)).

a vector of deterministic components of school preferences the *ex post* utility of student i from attending school j is

$$U_{ij} = \beta' Z_{ij} + \varepsilon_{ij} \quad (14)$$

To model strategic behaviours, we also parametrize $U_{ij}|p_{ij}$. In a simple behavioural model, when choosing the top choice school, parents may trade-off admission probability against other school attributes. This probability may enter the *ex-ante* utility additively. Alternatively, as suggested in [Chade and Smith \(2006\)](#), individuals may make choices based on expected utility such that the admission probability is multiplicative to the ex-post utility. To allow for *both* possibilities, we implement a flexible model that includes the admission probabilities as a standalone predictor of choices and also includes interactions of each observed characteristics with the these probabilities. **Note:** For ease of exposition, we model behaviours with respect to $q_{ij} = 1 - p_{ij}$ – i.e., the probability of rejection of person i from school j . It follows that

$$U_{ij}|p_{ij} = \beta_1' Z_{ij} + \beta_2'(Z_{ij}q_{ij}) + \pi q_{ij} + \varepsilon_{ij} \quad (15)$$

The parameter π captures the extent to which parents avoid schools that have low probability of admission (and therefore low expected utility), while β_2 captures the extent to which individuals “discount” particular school attributes with the admission probability. β_1 captures the relationship between choices and school attributes when the probability of rejection is zero (i.e., the probability of admission is 1). Importantly, this is the same parameter vector as in equation (14). Because we impose no restriction on the sign or magnitude of the interaction parameters (i.e., both β_2 and π can be positive, zero or negative), the model is sufficiently flexible to allow for individuals to be risk loving, risk neutral, or to exhibit varying degrees of risk aversion.

We further assume that ε_{ij} is distributed *i.i.d.* extreme value, that is $F(\varepsilon_{ij}) = e^{-e(-\varepsilon_{ij})}$. Under this standard distributional assumption (see [Train \(2009\)](#) and [McFadden \(1973\)](#)), and that of rational strategic behaviours, the probability that parent i submits a particular ranking over schools (i.e. Equation (13)) is simply a product of standard logit formulas.⁴³ Accordingly, where the parameter

⁴³That is, where q_{ij} is 1 minus the probability of admission for student i to school j , and parameter vector $\beta = [\beta_1, \beta_2, \pi]$, the probability that parent i chooses the ranking $\{r_{i1}, r_{i2}, \dots, R_i\}$ is:

$$\begin{aligned} Prob[r_{i1}, r_{i2}, \dots, R_i] = & \frac{\exp(\beta_1' Z_{ij}^{r_{i1}} + \beta_2'(Z_{ij}^{r_{i1}} q_{ij}^{r_{i1}}) + \pi q_{ij}^{r_{i1}})}{\exp(\beta_1' Z_{ij}^{r_{i1}} + \beta_2'(Z_{ij}^{r_{i1}} q_{ij}^{r_{i1}}) + \pi q_{ij}^{r_{i1}}) + \sum_{k=1}^{J-R_i} \exp(\beta_1' Z_{ik}^{r_{i0}} + \beta_2'(Z_{ik}^{r_{i0}} p_{ij}^{r_{i0}}) + \pi p_{ik}^{r_{i0}})} \\ & \cdot \frac{\exp(\beta_1' Z_{ij}^{r_{i1}})}{\exp(\beta_1' Z_{ij}^{r_{iR-1}})} \\ & \cdot \frac{\exp(\beta_1' Z_{ik}^{r_{ik}})}{\sum_{k=1}^{R_i} \exp(\beta_1' Z_{ik}^{r_{ik}})} \cdots \frac{\exp(\beta_1' Z_{ij}^{r_{iR-1}})}{\exp(\beta_1' Z_{ij}^{r_{iR-1}}) + \exp(\beta_1' Z_{ik}^{r_{iR-1}})} \end{aligned}$$

vector $\beta = [\beta_1, \beta_2, \pi]$, the log likelihood of observing all the choice lists for all parents is:

$$\log L(\beta) = \sum_{i=1}^N \log l_i(\beta) = \sum_{i=1}^N \log (\text{Prob}[r_{i1}, r_{i2}, \dots, R_i]). \quad (16)$$

Under the aforementioned behavioural, functional form, and distributional assumptions, one can estimate the relationship between school preferences and school attributes β_1 by estimating this model by maximum likelihood (i.e. finding the β vector that maximizes this expression).

Our model is conceptually similar to others in the literature but there are two key differences. First, we include an additional choice to the standard exploded logit model: In our first pseudo-observation, the individual chooses her first-ranked school over the set of *all* unranked schools in Trinidad and Tobago. As discussed above, including this additional first pseudo-observation allows us to anchor each individual’s choices to a common set of schools for all parents – making the choices and preferences comparable across individuals. The second key difference is that, as informed by the theory, when comparing the top choice to all unranked choices, we include the rejection probability as a covariate in the model, but we do not include it when comparing schools within the chosen list. These two modifications to the conventional multinomial logit model allow us to anchor each individual’s choice set while also explicitly accounting for strategic behaviors.⁴⁴

V.3 Estimating Admission Probabilities

When we compare the top choice school to all un-chosen schools, we account for the probability that student i would have been assigned to each school j had they applied. In many research settings, this probability is difficult to uncover. Fortunately, because we have many years of admissions data and students are assigned to schools based on a known algorithm, we can approximate this probability with the historical likelihood that student i would have scored above the cutoff for school j given their own incoming SEA score. We report these assignment probabilities for four different schools by the percentile of incoming SEA scores in Figure 4. School A is a very selective secondary school. Across all years, no student below the 82nd percentile scored above the cutoff for that school and all students above the 92nd percentile did. We show similar figures for less selective schools in other panels. Depending on their incoming SEA score, students may be marginal admits for some schools (predicted probabilities greater than zero and less than 1), be virtually guaranteed assignment at some schools, and have virtually no chance at others.

Because there is uncertainty for any student regarding their exact score, we compute the like-

⁴⁴Our model differs from [Hastings et al. \(2005\)](#) and [Hastings et al. \(2006\)](#) in that it uses a version of the exploded logit model with fixed coefficients, instead of estimating random coefficients by using mixed logit utility models. [Abdulkadiroğlu et al. \(2020\)](#) use the rank-ordered multinomial logit model to estimate a single measure of each school’s popularity separately for different covariate cells, whereas we use the modified version of the same model to estimate average population preferences for different school attributes.

likelihood that student i is within “striking distance” of a given cutoff as follows: For each school in each year and each SEA score, we code the “rough” likelihood as 0 if the score was more than 5 percentile points below the cutoff; 0.5 if it was within 5 percentiles of the cutoff; and 1 if it was above the cutoff by more than 5 percentiles. We then compute the probability of student i with a particular SEA score being assigned to school j (p_{ij}) as the average of these “rough” likelihoods for that SEA score across all years (excluding the year when the student actually applied). We also show these probabilities for the selected schools in [Figure 4](#). In our choice models, we use these rough probabilities and, to avoid extreme values, we truncate the rough probability at 0.05 and 0.95. So long as students are somewhat aware of these relationships based on historical precedent, our estimated probabilities proxy for the real admission probabilities used when making choices.

Indeed, in [Appendix E](#) we show that the admission probability strongly predicts school choices consistent with rational strategic behaviours – that is, conditional on observable measures of school quality, individuals are much less likely to rank schools at which their own rejection probability is high. More concretely, on average, a 20 percent higher likelihood of admission is associated with the same difference in choices as about a quarter of a standard deviation increase in incoming peer achievement or increasing the distance to school by 1.5 kilometers (just under a mile). While we emphasize the importance of modelling potential strategic behaviours, note that none of these modelling choices is consequential for our conclusions, as our main results are robust to excluding the admission probabilities all together.

V.4 Choice Parameter Estimates

We examine whether parents express preferences for schools based on their impacts on academic and non-academic dimensions, above and beyond easily observed school attributes. Our full estimation sample includes 329,481 households making school choice decisions. We estimate choice models separately for each (SEA score ventile) \times (gender) cell to allow preferences to vary based on the student’s gender and incoming achievement. Estimated standard errors are adjusted for clustering at the school-district level. Because the point estimates of the modified exploded multinomial logistic model are not easily interpretable, we report on the relative magnitudes and statistical significance of the estimated coefficients. We report on the coefficients on the standalone school attributes (i.e. β_1 from equation (15)) which represents the relationship between school preferences and choices when there are no strategic considerations – *i.e.*, when the probability of admission is one.⁴⁵ Except for the natural log of distance to school, all attributes have been standardized to be mean zero and unit variance (*for this section of the analysis only*).

We present results from two main models; (a) an Impacts Only Model (which includes schools’

⁴⁵Our main conclusions are the same when estimating more restrictive choice models that: (a) do not include neither the additional pseudo-observation nor account for admission probabilities (i.e., standard rank-ordered logit); and (b) include the additional pseudo-observation but do not account for admission probabilities ([See Appendix F](#)).

causal impact estimates for all outcomes, peer quality, and log distance), and (b) a Full Model that includes schools' causal impact estimates for all outcomes, the school-level averages for all the outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated "rough" likelihoods of school rejection (but only when comparing the top choice to all unranked schools) and interactions of these likelihoods with all school attributes. Because we seek to shed light on the relationship between schools causal effects and choices, we focus the discussion on these variables while others are included as controls.⁴⁶

To facilitate statistical inference, we pool estimates across different cells and report pooled coefficients and standard errors in Table 3.⁴⁷ Note that, for each model, we test for the significance of school impacts in six dimensions. To allow for a five percent type-I error among the six school impacts estimates, we use a simple Bonferonni adjustment. This requires an individual *alpha* of $0.05/6=0.0083$ or (with a two-sided alternative hypothesis) a *t*-statistic of 2.64. Because these are pooled averages across nonlinear models, while the reported *t*-statistics are valid for testing the null of zero relationships, the pooled averages should be interpreted with caution. As such, the coefficient estimates we use for interpretation come from the un-pooled models.

Before discussing the relationship between school impacts and school choices, we investigate the importance of proximity and incoming peer quality. The coefficient estimates for each cell are presented in Figure 5. With respect to distance, the figure reveals three key patterns. First, all students rank closer schools more highly - for all cells the estimate on distance is negative (with $p\text{-value}<0.01$) and the pooled models (Table 3) both have *t*-statistics over 20. Second, parents of the highest-achieving students (the top 20 percent) are somewhat more responsive to distance than those of lower-achieving students. Note that the patterns for the impacts only model (panel a) and the full model (panel b) are largely the same – indicating that distance to school is largely unrelated to other school attributes when shaping school preferences. Third, the relationship between choices and proximity for parents of boys and girls are similar within each achievement group.

A second key attribute when choosing a school is the academic achievement of peers (Hastings et al. 2005; Hastings et al. 2006; Hastings and Weinstein 2008). Figure 5 shows the coefficients on the potential peers' academic quality (the average SEA score of the incoming cohort). In both models, one rejects that choices are unrelated to peer achievement for every cell at the 5 percent level, and the pooled models (Table 3) both have *t*-statistics over 20. In the impacts only model (panel a), parents of higher-achieving boys are more responsive to peer quality than low-achieving counterparts. However, this difference largely goes away in the full model (panel b). Also, notice

⁴⁶See Appendix G for the estimated coefficients on the school-level averages outcomes from the Full Model.

⁴⁷We report the mean of the coefficient across cells. The standard error of this mean is computed assuming independence across cells.

that the magnitude of the estimated coefficients from the full model fall by over a third relative to the impacts only model – suggesting that the association between school preferences and better peers (higher incoming achievement) may partially reflect preferences for better average outcomes.

The importance of proximity and peer quality in shaping the schooling decision is consistent with other studies (e.g., [Hastings et al. 2005](#); [Abdulkadiroğlu et al. 2020](#)). A comparison of the coefficients on peer quality and proximity from the full models implies that, on average, increasing peer quality by about 0.445 standard deviations (roughly the difference between a student’s top choice school and the second choice school) is associated with the same difference in choices as doubling the distance between the primary and secondary school. That is, the choice parameters imply that parents may be willing to travel about 3.2 times farther to attend a secondary school with one standard deviation higher in incoming peer scores. The average distance is about 6 kilometers, so that this amounts to travelling an *additional* 13 kilometers (or 8 miles) to attend a school with one standard deviation higher peer achievement. This relative relationship is largely similar throughout the achievement distribution for both males and females.

V.4.1 Impacts on Academic Outcomes

The previous results suggest that parents in Trinidad and Tobago make similar choices to parents in other settings and appear to make rational choices. We now turn to the importance of schools’ causal impacts. [Figure 6](#) and [Figure 7](#) plot the estimates separately for each outcome. In the top panels, we plot coefficients on schools causal impacts from the impacts only model. In the bottom panels, we plot coefficients on schools causal impacts from the full model (that also includes school level averages). We focus on the relationship between choices and the school impacts, but we plot the coefficients on school level average outcomes from the full model in [Appendix G](#).

We first discuss the relationship between school preferences and school impacts on high-stakes exams. [Figure 6](#) (panel a) summarizes the impacts only results and reveals three patterns: (1) School impacts on high stakes exams are associated with the choices of most parents, (2) The choices of parents of high-achieving students are more strongly related to school impacts on high-stakes exams than those of parents of lower-achieving students, and (3) at all levels of incoming achievement, the choices of girls’ parents are more strongly related to school impacts on high-stakes exams than that of parents of boys. These patterns are illustrated by the positive and significant relationship between the individual’s score percentile and the coefficient magnitude (which is more pronounced for girls). In the full model, as we control for school averages (i.e., average exam scores and that for other outcomes), the relationship between school preferences and school impacts on high-stakes exams is appreciably weaker ([Figure 6](#), panel b) – indicating that some of the association between school impacts and parent preferences may have been driven by peer quality or average outcomes (as in [Abdulkadiroğlu et al. \(2020\)](#)). While one can reject that the choices of par-

ents of higher-achieving students are unrelated to school impacts on high-stakes (p -value <0.001), *in the full model* the choices of parents of children with low incoming scores are largely unrelated to school impacts on high-stake tests (indeed, the pooled t -statistics on high-stakes exam impacts are 9.1 and 0.14 for those above and below the median, respectively).⁴⁸ For those in the top 20 percent of incoming achievement, (taken at face value) the estimates imply that parents of girls and boys may be willing to travel roughly 30 and 15 kilometers farther to attend a secondary school at the 85th percentile of the high-stakes impact distribution than one at the median, respectively. This pattern for high-achieving students stands in contrast to [Abdulkadiroğlu et al. \(2020\)](#) who find that after conditioning on peer quality, no parent's school preferences are related to school impacts on school exit exam performance. We discuss possible reasons for these differences in Section [V.5](#).

Our next academic outcome is low-stakes exams. While there is strong evidence that certain parents may prefer schools that raise high-stakes exam performance, there is weak evidence that parents prefer schools that improve low-stakes exam performance (conditional on high-stakes performance).⁴⁹ Specifically, the impacts only model ([Figure 6](#), panel a) renders small insignificant estimates across most of the incoming achievement distribution for both boys and girls. In the full model the point estimates are larger in magnitude, but are mostly not statistically different from zero. Indeed, the pooled inference tests yields t -statistic of 2.05, which is below the threshold for rejection on 2.64 to account for multiple hypothesis testing. The figure does suggest that (in the full model only) parents of females choose schools with higher impacts on low-stakes exams while the opposite is true for boys. However, this pattern only exists in particular sub-samples of the data and is not robust across models (full and impacts only), so we take this as suggestive at best. In Trinidad and Tobago, average school outcomes on high-stakes exams are made public, while average school outcomes on low-stakes exams are not. As such, the stronger and more robust relationships between school preferences and high-stakes impacts than for low stakes impacts are consistent with (a) parents discerning school impacts on high-stakes exams but not on low-stakes exams, or (b) parents not caring about school impacts on low-stakes tests precisely because they are low stakes.

Our final academic outcome is dropout. As with low-stakes exams, the patterns suggest that parents are not systematically more likely to chose schools that causally reduce dropout. In the impacts only model, the relationship is not different from zero for most cells, but there is some

⁴⁸Looking at the school average high-stakes scores ([Appendix G](#)), the relationship between school choices and school averages mirror those of school's causal impacts - the choices of high-achieving males and females are associated with better school average high stakes performance, but not those of lower achieving males and females. These patterns are consistent with [Hastings et al. \(2006\)](#) who find that parents value schools with better average outcomes, or with [MacLeod and Urquiola \(2019\)](#) who argue that parents may value schools with better average outcomes if such attributes serve as positive signals in the labor market for example.

⁴⁹An insignificant or small estimated coefficient could indicate that either parents don't value that particular school attribute or, alternatively, that parents care about it but they don't have enough information about it. We favor the interpretation that an insignificant school feature does not play an important role in the schooling decision, remaining agnostic about which reason is more likely to occur in each particular case.

suggestive evidence that higher achieving parents may be more likely to chose schools that causally reduce dropout (Figure 6, panel a). In the full model, for none of the gender-by-achievement cells can one reject that the coefficient on dropout impacts is zero (Figure 6, panel b). Consistent with this, the pooled t -statistic in the full model is -0.411 so that one cannot reject that there is no relationship on average. As with low-stakes exam impacts, there are suggestive patterns for different samples, but these are generally not robust. As with the low-stakes exams, school-level dropout rates are not publicly reported. As such, the lack of a strong relationship between school impacts on dropout and school preferences may be because school impacts on dropout may be particularly difficult for parents to observe and therefore respond to.

V.4.2 Impacts on Non-Academic Outcomes

Next we document the relationship between parent school choices and school impacts on non-academic outcomes in Figure 7. Recall that all variables are coded so that positive values indicate better outcomes. We start with teen motherhood. The patterns from the impacts only model (panel a) reveal little association between school choices and impacts on teen motherhood for females, but a positive relationship for males. The pooled t -statistic is 2.27, so that (after accounting for multiple hypothesis testing), one does not reject that there is no association on average. However, in the full model that also includes the teen motherhood rate (and the averages for all other outcomes), the point estimates for females become positive at the lower end the distribution and many of the point estimates (at the bottom of the achievement distribution) for both males and females are positive and significantly different from zero at the 5 percent level (panel b). Consistent with this, in the full model the pooled t -statistic is 6.44, so that the null of no association on average is rejected. This relationship is driven primarily by those below the median on the distribution where this association is very strong (pooled t -statistic is 10.5).

Taken together, the results indicate that parents of low-achieving students (both males and female) choose schools that causally reduce teen motherhood. In the full model, the point estimates imply that the average parent of a male and female would be willing to increase their distance by about 6.2 and 1.3 kilometers to send their child to a school that was at the 85th percentile of the (non) teen motherhood effect distribution versus one at the median, respectively. These choice patterns are even stronger for parents of students in the bottom half of the incoming achievement distribution, where the average parent or a male and female would be willing to increase their distance by about 11 and 4.5 kilometers to send their child to a school that was at the 85th percentile

of the (non) teen motherhood effect distribution versus one at the median, respectively.⁵⁰ While it is true that only females can be mothers, both males and females are affected by teen pregnancies, so that a response from both males and females is not unreasonable. However, the stronger relationships for males is somewhat counterintuitive. This could reflect the fact that the choices of females may be more strongly related to school impacts on academic outcomes in general (as is the case), but we cannot rule out that school impacts on teen birth are correlated with some other school attribute that parents are responding to.

Another important non-academic measure is teen arrests. There is robust and strong evidence that the parent choices are related to school impacts on teen arrests. One can see this clearly in the impacts only model in [Figure 7](#) (panel a). Parents of boys and girls at all achievement levels are more likely to choose schools that reduce teen arrests. The figure reveals that one can reject zero association for most cells (the pooled t -statistic is 16.9), and that the relationship is stronger for parents of high-achieving students. In the full model (with averages included), the relationships are largely unchanged (panel b) – the point estimates all remain positive and the pooled t -statistic is 9.08. In sum, both boys and girls choose schools that causally reduce arrests. Overall, parents of the highest-achieving students appear to have the strongest preferences for schools that causally reduce crime. Point estimates suggest that parents of students in the bottom decile may be willing to increase their distance by about 3 kilometers to send their child to a school that was at the 85th percentile of the (non) teen arrest impact distribution versus one at the median, while that for those in the top decile is about 13 kilometers. These results are consistent with parents preferring schools with a lower prevalence of risky behaviors (conditional on peer achievement and average test score outcomes) and schools that reduce these behaviours (i.e., arrests) in their own children.

Given that most arrests are of males, the fact that parents of females are generally more responsive to school impacts on arrests than boys may seem odd. Similarly, because teen arrests are more prevalent among low-achieving students, the greater responsiveness among high-achieving students is not intuitive. We speak to these points in a number of ways. First, we point out that an aversion to crime *victimisation* would lead one to prefer schools that causally reduce crime even if one is not worried about one's own child committing a crime. As such, the reactions of high-achieving parents and parents of girls may reflect an aversion to crime victimisation. Second, we show that arrests do occur even among high achievers in [Appendix Figure B7](#), such that being arrested is not irrelevant for any group. Third, we point out that being arrested is an equilibrium outcome so

⁵⁰We also compare the implied distance a student would be willing to travel for a 1σ increase in high-stakes exams impacts vis-a-vis the distance a student would be willing to travel for a 1σ increase in (non) teen motherhood impacts. We find that low-achieving students may be willing to travel 5 to 10 kilometers farther to attend a school with 1σ higher (non) teen motherhood impact (reduce teen motherhood by 17.3 percentage points) than a school with 1σ higher high-stakes exam impact (raise scores by 0.44σ). For both, males and females this relation reverses as the incoming test scores improve ([Figure 8](#), left panel).

that parents that value low arrests *may* send their children to schools that reduce arrests (and also do other things to avoid arrests) leading them to have lower arrest rates in equilibrium. Finally, the pattern for arrests across groups may reflect a differential responsiveness to school attributes in *any* dimension across student types. To speak to this last point, we use the proximity estimates to compute the implied distance a student would be willing to travel for a 1σ increase in test score value-added compared to a 1σ increase in arrests value-added. Using this approach (which would account for general differences in responsiveness to school quality across student achievement levels and sex) the patterns are consistent with one's intuitions. The middle panel of Figure 8 shows that low-achieving males may be willing to travel 5 kilometers farther to attend a school with 1σ higher arrest value added (reduce arrests by 3.7 percentage points) than a school with 1σ higher high-stakes exam value added (raise scores by 0.44σ). In contrast, girls are similarly responsive to arrests and high stakes exam value-added. Looking at students at the top of the achievement distribution, the relationships are flipped. These high achieving students (both boys and girls) would travel farther for those same test score gains than for the reduction in arrests. In sum, when compared to the relationship between school preferences and high stakes exam impacts, parents of boys appear to place more relative weight on arrest value added than girls, and parents of low-achieving students (particularly boys) appear to place more relative weight on arrest value added than those of high-achieving students – patterns that align well with ones intuitions.

The last outcome we examine is formal employment at ages 27 and older. As with arrests, There is robust and strong evidence that all parents choose schools with positive causal impacts on employment (Figure 7). In the impacts only model (panel a), the pooled t -statistic is 13.8 - well above the multiple hypothesis adjusted threshold for rejecting no association. This is largely similar in the full model (panel b) – reinforcing the robustness of this relationship. In the full model, the pooled t -statistic is 10.26 - leading one to reject the null of no association. In the full model, the point estimates suggest that, on average, parents may be willing to increase their distance by about 5 kilometres to send their child to a school that was at the 85th percentile of the employment impacts distribution versus one at the median. These patterns are noteworthy for two reasons. First, this is the first direct demonstration that parents choose (i.e., prefer) schools that have causal impacts on formal employment (above and beyond peer quality and impacts on academic outcomes). Second, even though choices for children in the bottom half of the incoming achievement distribution are largely unrelated to school impacts on high-stakes test scores, they *are* related to school impacts on formal adult employment – suggesting that for more than half of the population, school impacts on employment may matter more than impacts on high-stakes exams.

To directly compare the relationship between high-stakes test-score value added and choices and those for formal labor market participation, the right panel of Figure 8 suggests that low-achieving males and females may be willing to travel 4 and 1 kilometers farther to attend a school

with 1σ higher labor market value added (increase employment by 7 percentage points) than a school with 1σ higher high stakes exam value added (raise scores by 0.44σ), respectively. In stark contrast, males and females at the top of the achievement distribution would travel about 3 and 10 kilometres farther for those same test score gains than for the increased employment. As with the arrests effects, the pattern of choices suggest that low-achieving students may place greater relative weight on school effects on non-academic outcomes (as compared to effects on academic outcomes) than high-achieving students for whom schools' test-score value added is very strongly associated with school preferences.

In sum, we show that parents choose, and therefore likely value, schools that have higher causal impacts on certain academic and non-academic outcomes. We show that this is not simply due to parents choosing schools with better average outcomes or better peers. Also, consistent with school quality being multidimensional, parents choose schools that have causal impacts on outcomes other than high-stakes tests such as crime and formal labor-market participation. Importantly, the correlations between school impacts on high stakes exams and impacts on arrests and formal labor market participation are relatively low. This suggests that strong parental preferences for school impacts on non-academic outcomes (that are largely unrelated to test score impacts) are a plausible explanation for the weak link between parental preferences and school impacts on test scores.⁵¹ It is important to point out that because school effects are very similar throughout the incoming test score distribution, these patterns are not due to schools having different effects on children by incoming achievement – a form of match effect.⁵² That is, we can rule out that our key results are driven by test score impacts having larger marginal effects for high-achievement children while non-academic impacts having larger effects on low-achieving children. Rather, our results likely reflect differences in preferences (or differences in information).⁵³

⁵¹Our models use the 2SLS estimated school impacts as explanatory variables across all years. Because the choice year is included when forming this estimate, one may worry about mechanical correlation between the estimated impacts and the desirability of the school. To assuage this concern, we estimate our choice models using leave-year-out 2SLS estimates. Because the 2SLS estimates are based on several instruments, leave-year-out estimates can vary a lot for the same school from year to year – introducing non-trivial estimation errors. As one might expect, (see [Appendix H](#)), our results using both leave-year-out estimated impacts and leave-year-out school average outcomes are noisier but qualitatively similar, and our central conclusions are unchanged.

⁵²In [Appendix Table B7](#), the maximum likelihood based correlations between the average effects and those at the 25th, 50th and 75th percentiles are all very high (above 0.88) for high stakes exams, low stakes exams, teen motherhood, and formal employment. For teen arrests, they are somewhat lower (above 0.72) but are still high.

⁵³**It is worth noting** that insofar as the equal-weighted school impacts are inaccurate measures of school impacts for particular kinds of students (i.e., there are large match effects in unobserved dimensions), it would bias our results toward zero – making it *less* likely to find any association between choices and estimated school impacts. As such, *even if there were considerable match effects in dimensions other than incoming achievement*, any systematic relationships we find between choices and school impacts would reflect real relationships.

V.5 Discussion of Parental Preference Results

One of our key findings is that parents may value school impacts on multiple outcomes above and beyond peer quality and average outcomes. The only other paper to formally test this notion is [Abdulkadiroğlu et al. \(2020\)](#) who find that parents do prefer schools that improve academic outcomes, but not after controlling for peer quality. Our results are a nice counterpoint to their work because we demonstrate that context matters. Also, by moving beyond academic outcomes and examining parental preferences for non-academic outcomes such as crime, teen fertility, and labor market participation, we shed light on the extent to which parents value school impacts beyond academics – this is very important given that many school choice evaluations use test scores alone.

Another potential explanation for differences between our findings and [Abdulkadiroğlu et al. \(2020\)](#) is market size. Several studies show that when individuals are faced with too many options they often opt for simplicity (e.g., [Iyengar and Kamenica 2010](#)), are more likely to rely on heuristics (e.g., [Besedeş et al. 2012](#)) and less likely to make the optimal choice (e.g., [Schram and Sonnemans 2011](#)). [Abdulkadiroğlu et al. \(2020\)](#) examine parent choices in the largest school district in the United States (which offers over 700 programs at over 400 schools). Their setting is a context in which sub-optimal behaviors are most likely to occur. In contrast, in our setting, individuals choose from a set of 134 options. While this is by no means a small market, it is much smaller than New York City (as are most markets), and therefore individuals' choices are less likely to be subject to errors induced by “overchoice.”

Our finding that school choices are related to school impacts on high-stakes examinations *only for parents of high-achieving students* relates to the overall lack of robust achievement effects, on average, of attending schools that parents prefer ([Beuermann and Jackson, 2020](#)). However, in the Trinidad and Tobago context, school impacts may be easier to infer for relatively sophisticated parents. Average incoming scores are well known and publicly reported. Additionally, school averages for the high-stakes exams are also reported at the school level. As such, it is plausible for a relatively sophisticated parent to observe schools with similar average outcomes and infer which one likely has larger impacts (based on average incoming test scores). In settings where average incoming scores are not reported or well known, this calculation may be much more difficult to conduct – offering another plausible explanation for our finding that the preferences of some parents (i.e., those of higher achieving students) are related to school's test score impacts (conditional on average outcomes) while some other studies do not find so. However, the fact that parents' school preferences are systematically related to schools' causal impacts on arrests and employment (even conditional on school averages and peer quality) suggests that some parents may be able to infer school impacts even when information is imperfect (perhaps through some combination of knowing the incoming student characteristics and reputation effects regarding average outcomes).

VI Conclusions and Policy Implications

Individual schools have meaningful causal effects on an array of outcomes; these include low-stakes test scores, dropout, teen motherhood, high-stakes school leaving exams, being arrested, and formal labor market participation. However, consistent with school quality being multidimensional, the correlations between school impacts on high-stakes tests and other outcomes is surprisingly low. From a policy perspective, our results suggest that school impacts on test scores may not be the best measure of a school's impacts on longer-run outcomes. Accordingly, policymakers should be cautious (and thoughtful) regarding using test score impacts in accountability systems and incentive pay schemes and may wish to adopt a more holistic view of school quality.

Linking causal school impacts to choice data, we find that parents choose schools that have larger positive impacts on high-stakes tests and *also* those that decrease crime and increase labor market participation. These patterns persist even conditional on average school outcomes and peer quality. These results suggest that parents may use reasonable measures of school quality when making investment decisions for their children – a requirement for the potential benefits of school choice (Friedman, 1955). The fact that parents do not *only* choose schools that improve academics but also those that improve non-academic and longer-run outcomes suggests that the benefits to school choice may extend to a wide range of outcomes (not just test scores). This result provides a plausible explanation for the fact that parental preferences for schools are not strongly related to school's test score impacts (MacLeod and Urquiola, 2019). It also suggests that policy evaluations based solely on test scores may be misleading about the effects of school choice on welfare.

We find important heterogeneity in parent choices. High-achieving students' choices are more strongly related to schools estimated impacts on high-stakes exams than impacts on non-academic outcomes, while the choices of low-achieving students' are more strongly related to school's impacts on non-academic outcomes than those on high-stakes exams. This suggests that market forces may drive competition more strongly to raise test scores among schools serving high-achieving populations and non-academic outcomes among schools serving low-achieving populations. If these differences reflect parents' true preferences, this may be efficient. However, if these differences across parents reflect differences in information, there may be value to the provision of information to parents regarding the causal impacts of schools (as opposed to school averages) on a wide array of both academic and nonacademic outcomes.⁵⁴ The provision of such information may improve the decisions of all parents and could increase the *potential* allocative efficiencies and competitive benefits of school choice.

⁵⁴A recent experimental study in Chile (Allende et al. 2019) find that a personalized information intervention led parents to chose schools with higher causal impacts which in turn improved their children's academic achievement.

Tables and Figures

Table 1: Summary Statistics

	All Schools	Male	Female	Above median	Below median
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: SEA data (cohorts: 1995 - 2012)</i>					
Female (%)	50.46 (50.00)			53.99 (49.84)	46.94 (49.91)
Admitted cohort size	231.17 (177.65)	233.75 (176.52)	228.78 (178.65)	168.15 (130.07)	325.16 (196.56)
Standardized SEA score	0.00 (1.00)	-0.13 (1.04)	0.13 (0.94)	0.70 (0.64)	-0.70 (0.78)
Individuals	329,481	163,217	166,264	164,519	164,962
<i>Panel B: NCSE data (linked to SEA cohorts: 2006 - 2012)</i>					
Took NCSE (%)	90.22 (29.70)	88.34 (32.10)	92.10 (26.98)	95.18 (21.42)	82.54 (37.96)
Standardized NCSE score	0.00 (1.00)	-0.22 (1.00)	0.21 (0.96)	0.30 (0.90)	-0.55 (0.94)
Individuals	111,294	55,517	55,777	67,616	43,678
<i>Panel C: CSEC data (linked to SEA cohorts: 1995 - 2011)</i>					
Took at least 1 subject (%)	73.22 (44.28)	66.96 (47.04)	79.36 (40.47)	86.98 (33.66)	59.78 (49.03)
Number of subjects passed	3.17 (3.09)	2.55 (2.95)	3.78 (3.11)	4.83 (2.96)	1.55 (2.25)
Qualified for tertiary (%) *	34.55 (47.55)	27.53 (44.66)	41.45 (49.26)	58.05 (49.35)	11.61 (32.04)
Individuals	313,580	155,322	158,258	154,921	158,659
<i>Panel D: CAPE data (linked to SEA cohorts: 1999 - 2009)</i>					
Took at least 1 unit (%)	19.82 (39.86)	15.33 (36.03)	24.24 (42.85)	33.59 (47.23)	4.35 (20.40)
Number of units passed	1.40 (2.93)	1.06 (2.61)	1.73 (3.18)	2.40 (3.54)	0.27 (1.33)
Earned Associate Degree (%)	14.73 (35.44)	10.87 (31.12)	18.54 (38.87)	25.78 (43.74)	2.32 (15.05)
Earned scholarship (%)	0.95 (9.68)	0.66 (8.11)	1.23 (11.01)	1.78 (13.23)	0.01 (0.84)
Individuals	208,794	103,682	105,112	110,444	98,350
<i>Panel E: Criminal records (linked to SEA cohorts: 1995 - 2010) - in percent</i>					
Arrested by 18	3.27 (17.79)	5.78 (23.34)	0.81 (8.97)	1.75 (13.10)	4.72 (21.22)
Individuals	297,948	147,544	150,404	145,288	152,660
<i>Panel F: Birth records (linked to SEA cohorts: 2004 - 2010) - in percent</i>					
Live birth by 19			10.11 (30.15)	6.86 (25.27)	15.19 (35.90)
Individuals			43,834	26,715	17,119
<i>Panel G: Labor market data (linked to SEA cohorts: 1995 - 2006) - in percent</i>					
Formally employed	75.52 42.99	79.38 40.45	71.77 45.01	78.28 41.24	73.64 44.06
Individuals	160,912	79,319	81,593	65,420	95,492

Notes: Standard deviations reported in parentheses below the means. *Qualification for tertiary education requires passing five CSEC examinations including English language and mathematics. Columns (4) and (5) report statistics differentiated by the rank of the assigned school based on the SEA score mean of students assigned to each school.

Table 2: Standard Deviations and Correlations of Persistent School Impacts

Outcome	School Level ($\sigma_{\theta^{TOR}}$)	School Level Correlations with High-Stakes	
	Size of Impact	Average	75th %ile of the achievement distribution
	(1)	(2)	(3)
Standardized outcomes			
High-Stakes Index	0.441 [0.397 , 0.483]	1.000	1.000
Low-Stakes Index	0.473 [0.433 , 0.508]	0.100 [-0.042 , 0.224]	0.032 [-0.106 , 0.177]
Binary outcomes			
No Dropout by 14	0.090 [0.077 , 0.100]	0.121 [0.019 , 0.228]	0.117 [-0.044 , 0.290]
Not arrested by 18	0.037 [0.032 , 0.041]	0.282 [0.164 , 0.419]	0.407 [0.210 , 0.619]
No live birth by 19	0.173 [0.147 , 0.193]	-0.036 [-0.171 , 0.087]	0.174 [0.021 , 0.311]
Formally employed 27+	0.070 [0.058 , 0.079]	0.152 [0.025 , 0.294]	0.014 [-0.148 , 0.157]

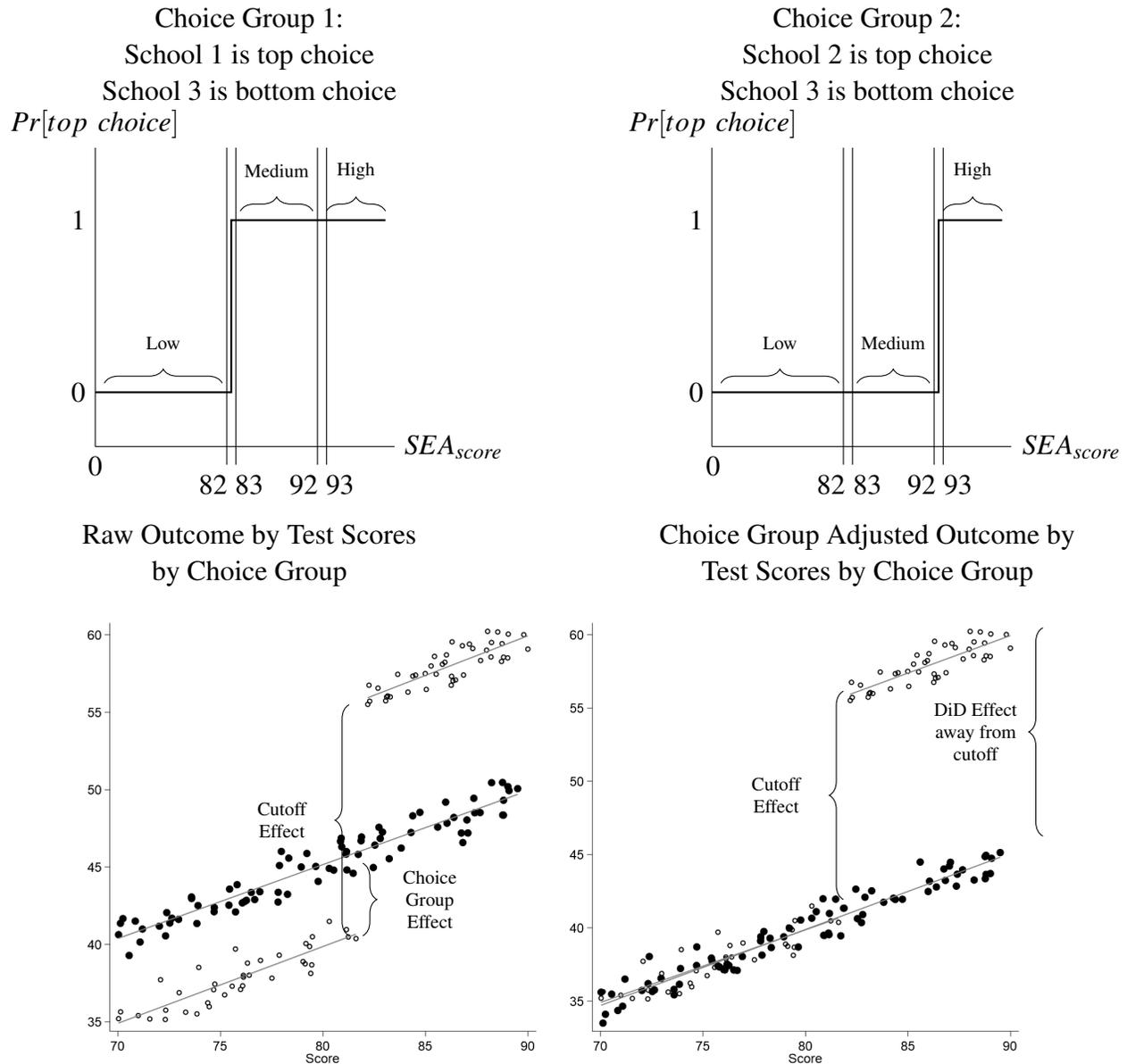
Notes: All estimates shown were computed by bootstrap with 1,000 repetitions of the maximum likelihood approach described in the text. We report the median as the point estimate, as well as the 5th and 95th percentiles for the confidence intervals. Column (1) reports estimated standard deviations of the persistent school impacts for each outcome. Column (2) reports estimated correlations of the persistent school impacts on the high-stakes index (for the average student) with the persistent school impacts on other outcomes (also for the average student). Column (3) reports estimated correlations of the persistent school impacts on the high-stakes index (estimated with weights centered around the 75th percentile of the achievement distribution) with the persistent school impacts on other outcomes (estimated with weights centered around the 75th percentile of the achievement distribution). We do this using $weight_i = (1 + \frac{(75-pct_i)^2}{100})^{-1}$, where pct_i is the student's percentile in the incoming achievement distribution. We removed schools with outlier estimated impacts (i.e. beyond 4σ of the median school).

Table 3: Inference Tests

	Impacts Only Model			Full Model					
	Pooled average	Above median score	Below median score	Pooled average	Above median score	Below median score	Female	Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Distance	-0.236 (0.008)	-0.261 (0.013)	-0.211 (0.010)	-0.245 (0.008)	-0.275 (0.012)	-0.215 (0.010)	-0.226 (0.011)	-0.264 (0.011)	
Peer quality	0.906 (0.022)	1.101 (0.041)	0.711 (0.015)	0.555 (0.019)	0.598 (0.034)	0.512 (0.018)	0.503 (0.021)	0.607 (0.032)	
<i>Standardized value-added on</i>									
High-Stakes Index	0.620 (0.013)	0.984 (0.022)	0.255 (0.014)	0.217 (0.030)	0.440 (0.048)	-0.005 (0.035)	0.331 (0.047)	0.103 (0.035)	
Low-Stakes Index	0.031 (0.009)	0.033 (0.013)	0.029 (0.011)	0.039 (0.019)	0.042 (0.034)	0.035 (0.016)	0.177 (0.033)	-0.100 (0.019)	
No Dropout by Age 14	0.054 (0.019)	0.119 (0.032)	-0.011 (0.020)	-0.007 (0.017)	0.045 (0.029)	-0.058 (0.016)	0.030 (0.020)	-0.043 (0.027)	
No Live Birth by Age 19	0.041 (0.018)	0.064 (0.030)	0.019 (0.020)	0.116 (0.018)	0.044 (0.032)	0.189 (0.018)	0.045 (0.024)	0.188 (0.028)	
Not Arrested by Age 18	0.390 (0.023)	0.532 (0.034)	0.247 (0.031)	0.227 (0.025)	0.337 (0.039)	0.118 (0.031)	0.276 (0.032)	0.178 (0.038)	
Formally Employed at Age 27+	0.180 (0.013)	0.264 (0.022)	0.096 (0.014)	0.154 (0.015)	0.240 (0.026)	0.068 (0.014)	0.157 (0.016)	0.151 (0.024)	

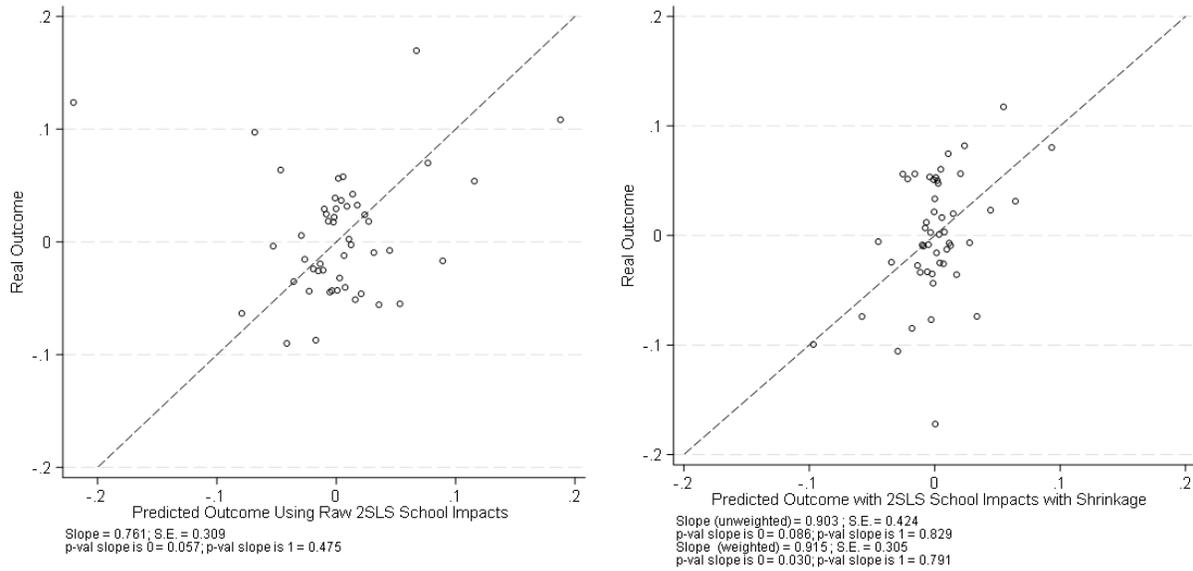
Notes: This table presents pooled averages of the choice model estimates for log distance to the school, peer quality, and standardized value-added school attributes (i.e., β_1 from equation 15) for two main models: (a) an Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance (presenting the aggregated pooled average (Column 1) and the average by incoming achievement level (Columns 2 and 3)) and (b) a Full Model that includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance (presenting the aggregated overall average (Column 4), and the averages by incoming achievement level (Columns 5 and 6), and by gender (Columns 7 and 8)). All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The pooled means are computed by taking the average across the point-estimates of all cells in a subgroup. Assuming independence across cells, the pooled standard error is computed by summing all subgroup variances, dividing by the number of cells in the subgroup, and taking the squared-root of the ratio. The pooled standard errors are presented in parentheses below each pooled average estimate.

Figure 1: Exemplar of Variation



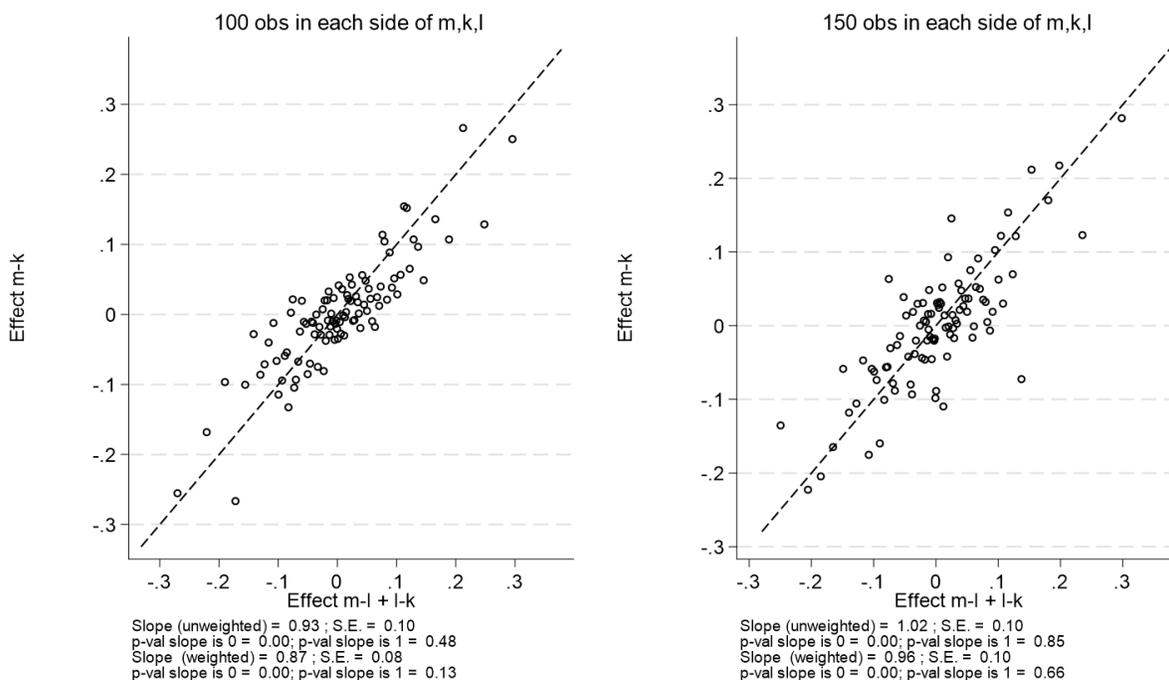
Notes: In the top panels, the Y-axis represents the probability of student i being assigned to her top choice; while the X-axis represents the student's SEA score. The top left panel shows the cutoff for school 1 and the top right panel shows the cutoff for school 2. In both panels those who score below the cutoff for the preferred school are assigned to school 3. This figure illustrates the two different sources of variation. **The RD variation** identifies the effect of being assigned to school 1 (with respect to school 3) by comparing the outcomes of the Low scoring group (hollow circles right below the cutoff for school 1 in the bottom left panel) to the Medium scoring group (hollow circles right above the cutoff for school 1 in the bottom left panel). This is labeled as "Cutoff Effect". **The Difference in Difference variation** comes from making comparisons *across* cutoffs even among students who are away from the cutoff. For example, the difference in outcomes between the Low scoring group at Choice Group 1 (hollow circles in bottom left panel) and the Low scoring group at Choice Group 2 (solid circles in bottom left panel) will reflect differences in choices (as both groups were assigned to the same school 3) – labeled as "Choice Group Effect". Therefore, provided that the effects of choices and test scores on outcomes are additively separable, the Cutoff Effect of attending school 1 versus 3 can also be identified exploiting observations away from the cutoff after Choice Group Effects are accounted for (graphically shown in the bottom right panel).

Figure 2: Predicted Cutoff Effects versus Actual Cutoff Effects



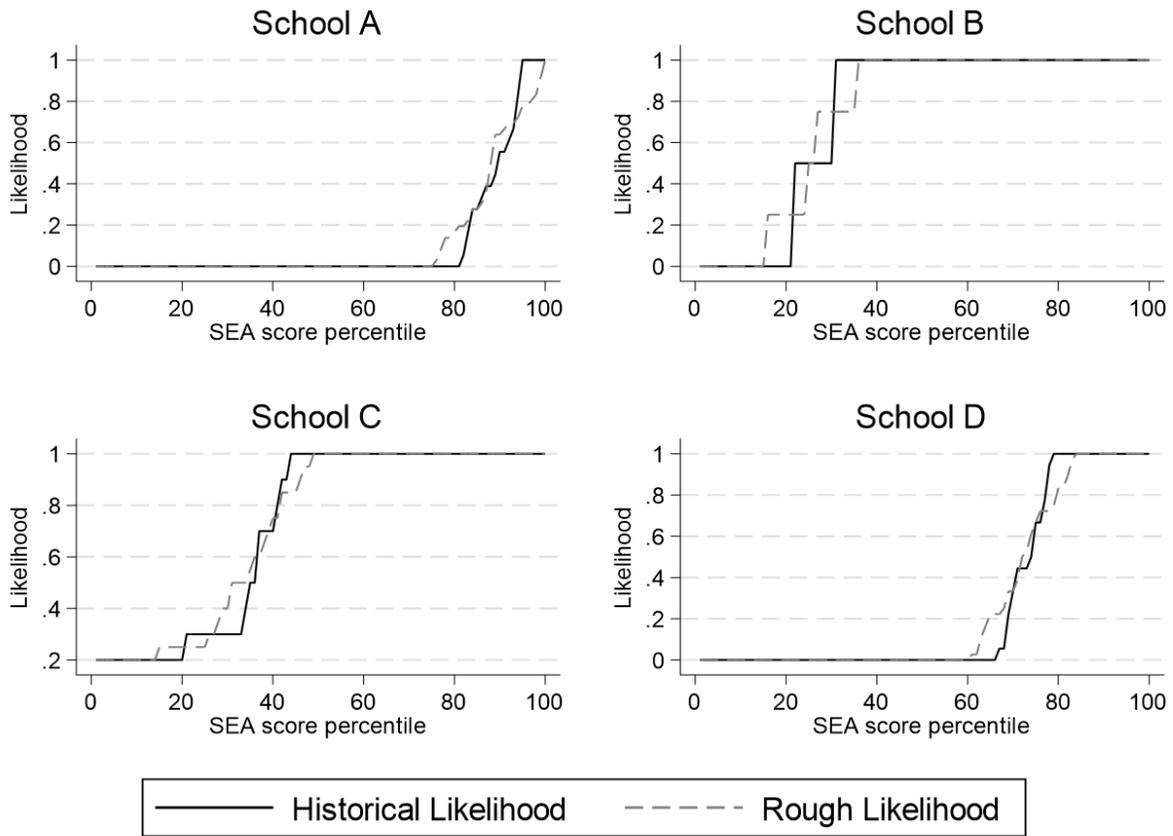
Notes: In both panels, the X-axis represents the estimated coefficients on an indicator for scoring above the rule-based cutoff resulting from an RD model that controls for a fifth degree polynomial of the SEA score, gender, district of residence at SEA registration, and religion; estimated for each school j and for each outcome where the estimated TOT school impacts ($\hat{\theta}_j^{TOTiv}$) enter as dependent variables. The Y-axis represents the estimated coefficients on an indicator for scoring above the rule-based cutoff resulting from an RD model that controls for a fifth degree polynomial of the SEA score, gender, district of residence at SEA registration, and religion; estimated for each school j and for each outcome where the individual level outcomes enter as dependent variables. Estimated slope and p-values resulting from testing for whether the slope differs from both 0 and 1 are shown below the graph. The short dashed line shows the 45° relationship. Schools have been grouped in 50 bins across the X-axis. Outliers above 4 standard deviations away from the median were removed. In the left panel, we show unweighted results. In the right panel, all school estimated effects are adjusted by the reliability ratio described in the text. We show unweighted results and results were all school estimated effects are weighted by the inverse of the squared standard error of each estimated coefficient for the real outcome.

Figure 3: Additivity of the School Effects.



Notes: Both panels display estimated relative school effects pooled across all outcomes. The left panel shows results where the estimated effects were obtained with least 100 students assigned to m , l or k . The right panel restricts this to at least 150 students. We show slopes and standard errors for the unweighted relationship and weighted by the sum of observations used to estimate each pair $m - l$ and $l - k$. The p -values associated with the hypothesis that the slope is 0 and that the slope is 1 are shown below these for each estimation. The short dashed line shows the 45° relationship. We use the mean of the estimated standard errors of the school effects when computing the reliability ratio. Similar results are obtained when employing the median of the standard errors. In this case, slopes are slightly lower, in average 0.03 units less than when using the mean and we still fail to reject all slopes being different from 1 at the 5% level. Observations have been grouped in 100 bins across the X-axis.

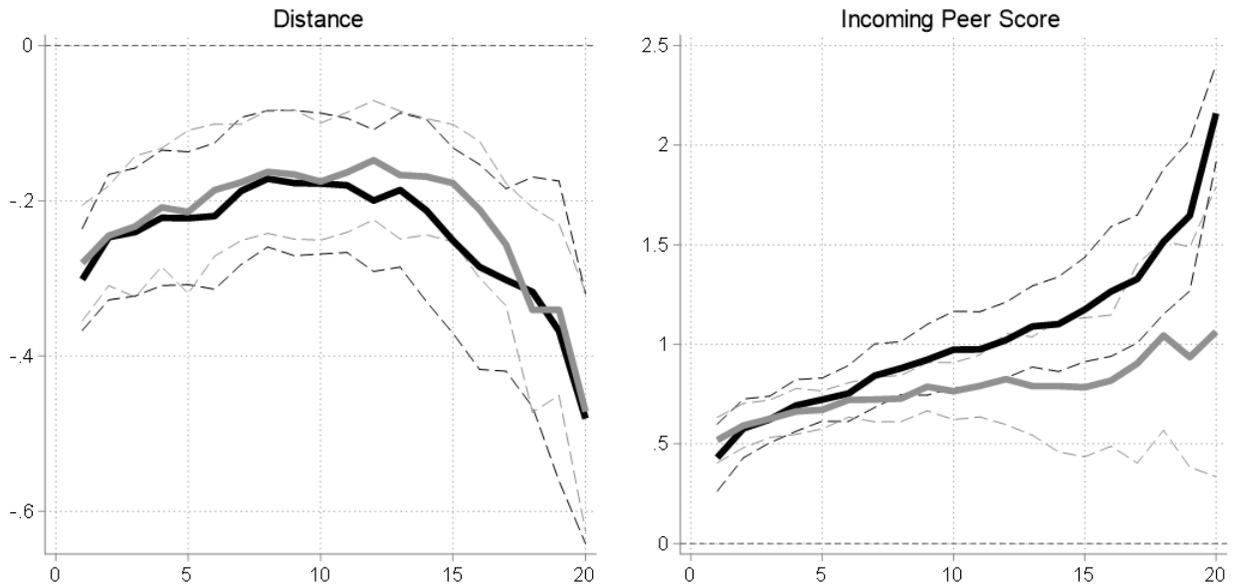
Figure 4: Estimated Admission Probabilities for Selected Schools



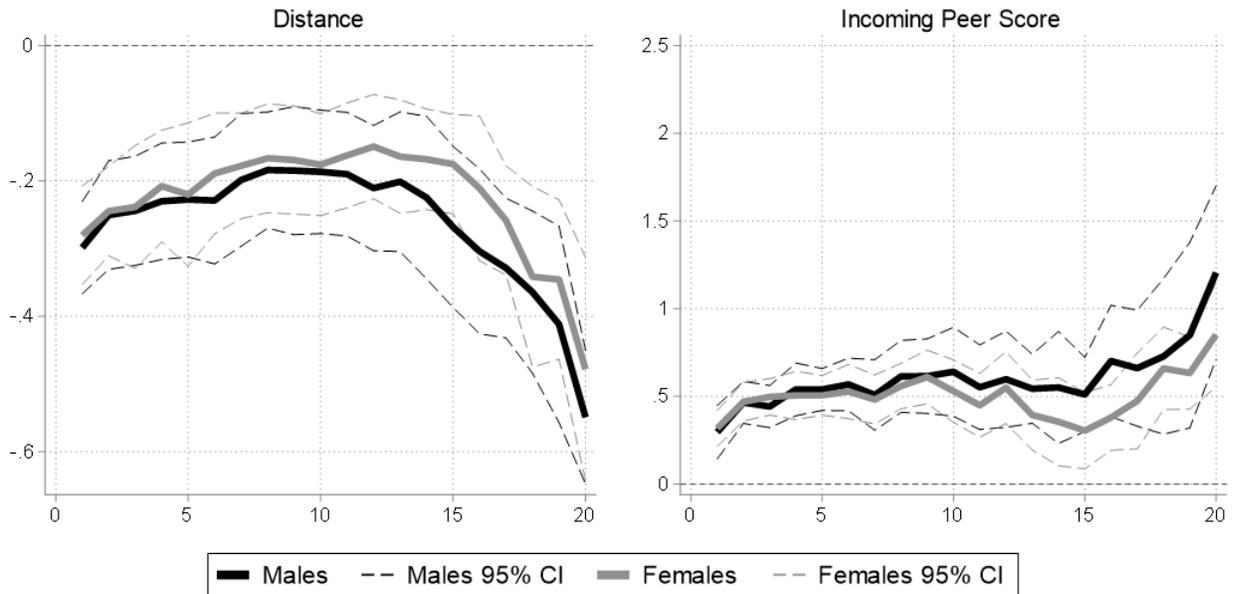
Notes: The X-axis represents the SEA score percentile. The Y-axis represents the likelihood of being assigned to a given school. School A is a very selective secondary school, where the school algorithm-based assignment cutoff has always been above the 82nd percentile. School B, on the other hand, is a less selective school where students above the 30th percentile have always scored above the cutoff. Historical probabilities are depicted with black hollow circles. “Rough” probabilities (calculated as described in the text) are depicted with gray diamonds.

Figure 5: Proximity and Incoming Peer Quality

(a) Impacts Only Model



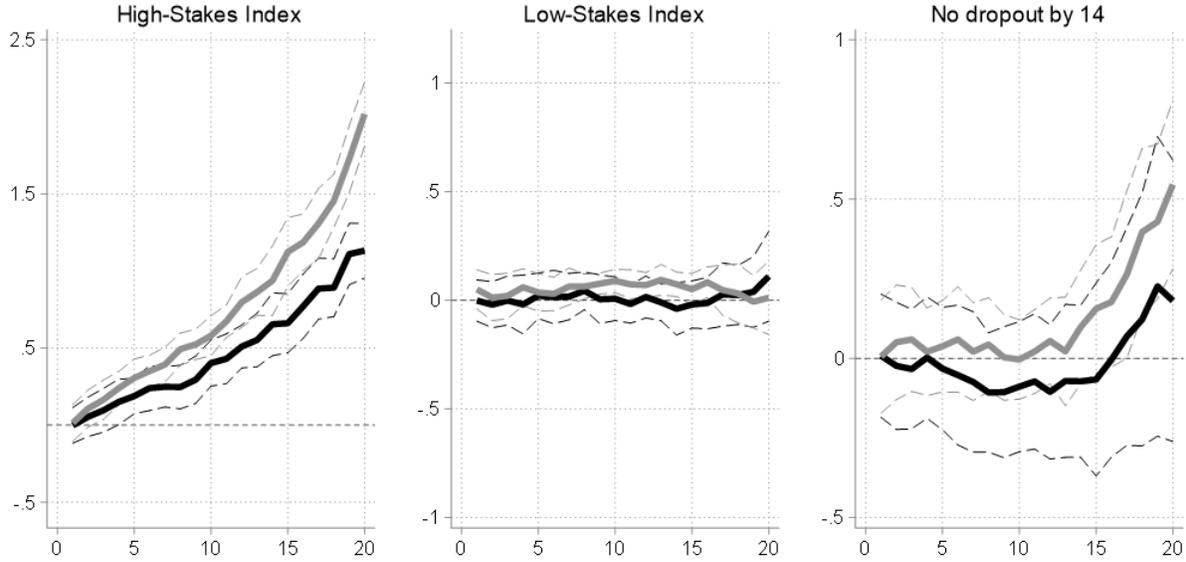
(b) Full Model



Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Figure 6: Academic Outcomes

(a) Causal Impact (Impacts Only Model)



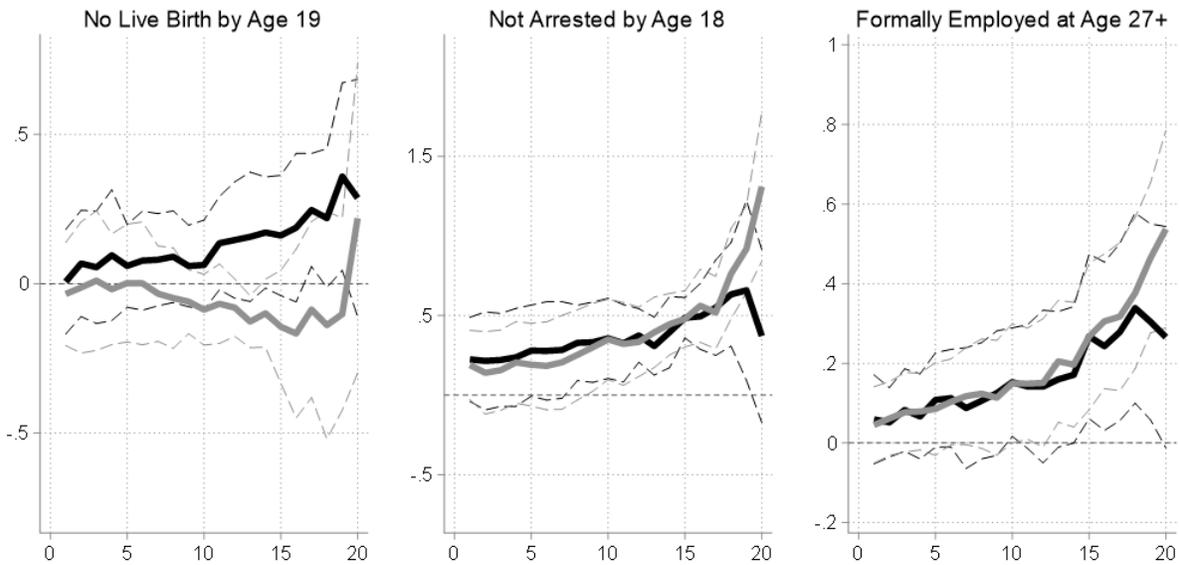
(b) Causal Impact (Full Model)



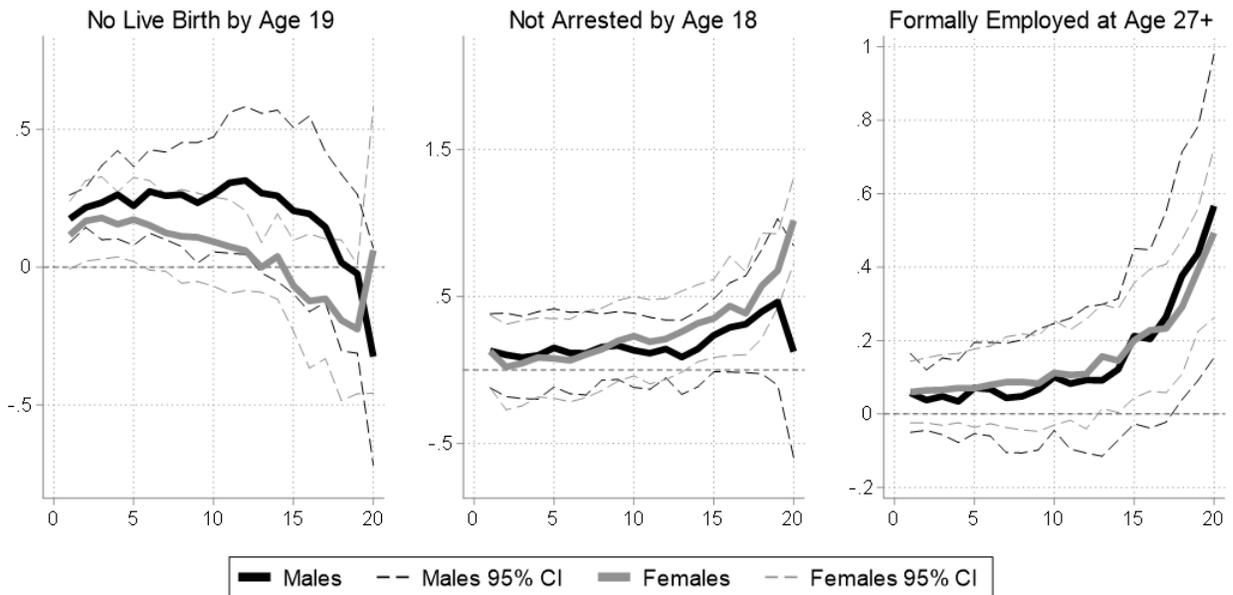
Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Figure 7: Non-Academic Outcomes

(a) Causal Impact (Impacts Only Model)

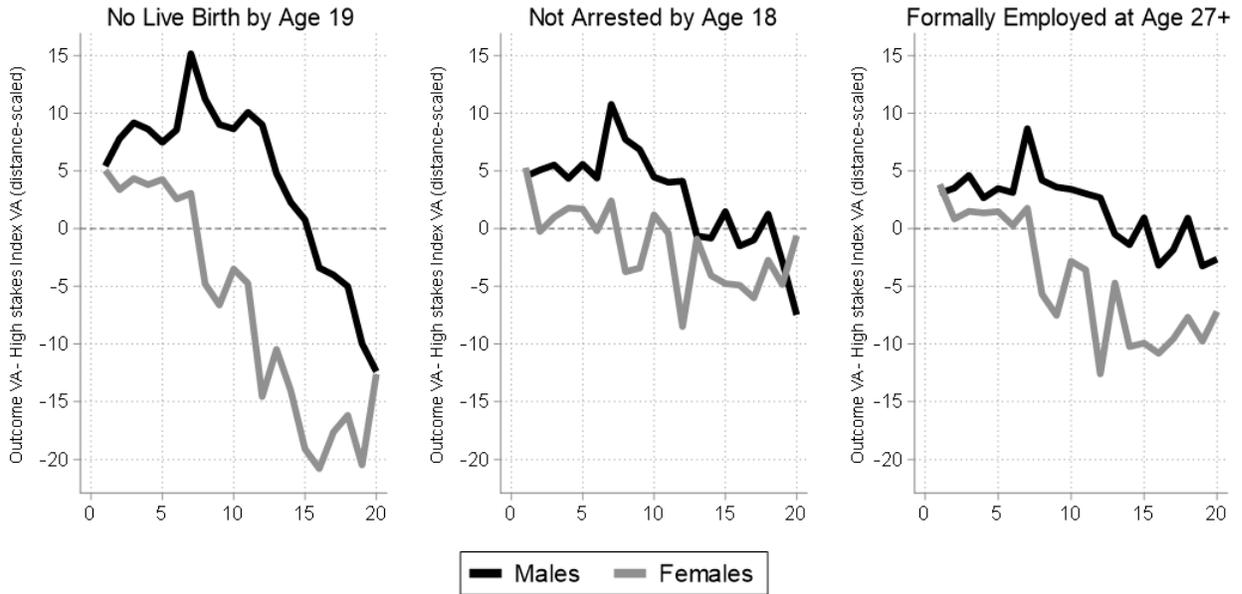


(b) Causal Impact (Full Model)



Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Figure 8: Relative Comparisons of Choice Model's Estimated Coefficients



Notes: This figure presents the difference between the choice model's estimates on the school impacts of three non-academic outcomes and the choice model's estimate on the school impacts of the high-stakes index, scaled by the log distance estimate. The X-axis represents the individual score ventile. The connected lines represent the difference between the choice model estimate on the non-academic impacts and the choice model estimate on the high-stakes impacts divided by the log distance cell estimate (and scaled by 6 for ease of interpretation). This difference is computed separately for each (SEA score ventile)×(gender) cell. The estimates result from the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality and log distance, control variables (whether the secondary school is on the same island, whether it is all-girls, whether it is all boys), the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes.

Acknowledgments. We are deeply grateful to Sabine Rieble-Aubourg and Dana King from the Inter-American Development Bank (IADB) for their critical support in establishing the necessary contacts to assemble the data used in the study. We are indebted to Therese Turner-Jones, Carina Cockburn, Musheer Kamau, David Rosenblatt, Inder Ruprah, and Moisés Schwartz from the IADB for their invaluable support during the elaboration of this project. We deeply thank Harrilal Secharan, Ria Boafo, Lisa Henry-David, Shalini Maharaj, Brenda Moore, and Peter Smith from the Trinidad and Tobago Ministry of Education for facilitating access to the educational data needed for the study, their assistance, and their generosity. We thank Registrar General of Trinidad and Tobago Karen Bridgewater for kindly granting us access to the national birth records; Amos Sylvester from the Crime and Problem Analysis Branch of the Trinidad and Tobago Police Service for facilitating access to arrests records; and Executive Director Niala Persad of the National Insurance Board of Trinidad and Tobago, as well as Andy Edwards, Arlene Grant, Feyaad Khan and Bernard Smith for their support and generosity while working in their facilities to match employment records while maintaining individual confidentiality. Tatiana Zárate and Diego Zúñiga provided excellent research assistance. This paper benefited from comments by Joshua Angrist, Samuel Berlinski, Matias Busso, Julian Cristia, Veronica Frisancho, Ofer Malamud, Norbert Schady, Diego Vera-Cossio, four anonymous referees, and seminar participants at the NBER, MIT, U. Chicago, LACEA, IADB, Zurich, Lund, and Wharton.

References

- Atila Abdulkadiroğlu, Joshua Angrist, and Parag Pathak. The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196, 2014.
- Atila Abdulkadiroğlu, Joshua D Angrist, Yusuke Narita, Parag A Pathak, Don Andrews, Tim Armstrong, Eduardo Azevedo, Yeon-Koo Che, Glenn Ellison, Brigham Frandsen, John Friedman, Justine Hastings, Guido Imbens, Jacob Leshno, Whitney Newey, Ariel Pakes, and Pedro Sant’anna. Breaking Ties: Regression Discontinuity Design Meets Market Design. Discussion Paper 2170, Cowles Foundation for Research in Economics. Yale University., 2019.
- Atila Abdulkadiroğlu, Parag Pathak, Jonathan Schellenberg, and Christopher Walters. Do Parents Value School Effectiveness? *American Economic Review*, 110(5):1502–1539, 2020.
- Nikhil Agarwal and Paulo Somaini. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86:391–444, 3 2018. ISSN 1468-0262. doi: 10.3982/ECTA13615. URL <https://onlinelibrary.wiley.com/doi/full/10.3982/ECTA13615><https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA13615><https://onlinelibrary.wiley.com/doi/10.3982/ECTA13615>.
- Claudia Allende, Francisco Gallego, and Christopher Neilson. Approximating the Equilibrium Effects of Informed School Choice. *Princeton University Mimeo*, 2019.
- Joshua Angrist, Peter Hull, Parag A. Pathak, and Christopher R. Walters. Simple and credible value-added estimation using centralized school assignment. *NBER Working Paper 28241*, 12 2020. doi: 10.3386/W28241. URL <https://www.nber.org/papers/w28241>.
- Joshua Angrist, Peter Hull, Parag A Pathak, and Christopher Walters. Credible school value-added with undersubscribed school lotteries *. *NBER Conference Paper*, 2021. URL <https://www.ccas-project.org/>.

- Joshua D Angrist, Parag A Pathak, and Christopher R Walters. Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics*, 5(4):1–27, 2013.
- Joshua D. Angrist, Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak, and Christopher R. Walters. Stand and Deliver: Effects of Boston’s Charter High Schools on College Preparation, Entry, and Choice. *Journal of Labor Economics*, 34(2):275–318, 2016.
- Christopher N. Avery, Mark E. Glickman, Caroline M. Hoxby, and Andrew Metrick. A Revealed Preference Ranking of U.S. Colleges and Universities. *The Quarterly Journal of Economics*, 128(1):425–467, 2013.
- Tibor Besedeş, Cary Deck, Sudipta Sarangi, and Mikhael Shor. Age effects and heuristics in decision making. *Review of Economics and Statistics*, 94(2):580–595, 2012.
- Diether W. Beuermann and C. Kirabo Jackson. The Short and Long-Run Effects of Attending The Schools that Parents Prefer. *Journal of Human Resources*, pages 1019–1053R1, apr 2020.
- Kevin Booker, Tim R. Sass, Brian Gill, and Ron Zimmer. The Effects of Charter High Schools on Educational Attainment. *Journal of Labor Economics*, 29(2):377–415, 2011.
- Simon Burgess, Ellen Greaves, Anna Vignoles, and Deborah Wilson. What Parents Want: School Preferences and School Choice. *The Economic Journal*, 125(587):1262–1289, 2015.
- Matias D. Cattaneo, Luke Keele, Rocío Titiunik, and Gonzalo Vazquez-Bare. Extrapolating treatment effects in multi-cutoff regression discontinuity designs. <https://doi.org/10.1080/01621459.2020.1751646>, 2020. doi: 10.1080/01621459.2020.1751646. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.2020.1751646>.
- Hector Chade and Lones Smith. Simultaneous Search. *Econometrica*, 74(5):1293–1307, 2006.
- Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9):2593–2632, 2014.
- John E Chubb and Terry M. Moe. *Politics, markets, and America’s schools*. 1990.
- David J. Deming. Better Schools, Less Crime? *The Quarterly Journal of Economics*, 126(4): 2063–2115, 2011.
- David J. Deming. Using School Choice Lotteries to Test Measures of School Effectiveness. *American Economic Review*, 104(5):406–411, 2014.
- David J. Deming, Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. School Choice, School Quality, and Postsecondary Attainment. *The American Economic Review*, 104:991–1013, 2014.
- Jonah Deutsch, Brian Gill, and Matthew Johnson. The Promotion Power Impacts of Louisiana High Schools (Executive Summary). *Mathematica Policy Research Reports*, 2020. URL <https://www.mathematica.org/our-publications-and-findings/publications/es-the-promotion-power-impacts-of-louisiana-high-schools>.
- Will Dobbie and Roland Fryer. Charter Schools and Labor Market Outcomes. *Journal of Labor Economics*, 38(8), 2020.
- Will Dobbie and Roland G. Fryer. The Medium-Term Impacts of High-Achieving Charter Schools. *Journal of Political Economy*, 123(5):985–1037, 2015.
- Milton Friedman. *The Role of Government in Education*. University of Chicago Press, 1955.

- D. Gale and LS Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.
- Guillaume Haeringer and Flip Klijn. Constrained school choice. *Journal of Economic Theory*, 144(5):1921–1947, 2009.
- Eric A. Hanushek. Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review*, 61(2):280–288, 1971.
- Justine Hastings, Thomas Kane, and Douglas Staiger. Parental Preferences and School Competition: Evidence from a Public School Choice Program. NBER Working Paper 11805, Cambridge, MA, 2005.
- Justine Hastings, Thomas Kane, and Douglas Staiger. Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery. NBER Working Paper 12145, Cambridge, MA, 2006.
- Justine Hastings, Christopher Neilson, and Seth Zimmerman. The Effects of Earnings Disclosure on College Enrollment Decisions. NBER Working Paper 21300, Cambridge, MA, 2015.
- Justine S. Hastings and Jeffrey M. Weinstein. Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *Quarterly Journal of Economics*, 123(4):1373–1414, 2008.
- Justine S Hastings, Thomas J Kane, and Douglas O Staiger. Heterogeneous Preferences and the Efficacy of Public School Choice. *Working Paper*, 2009.
- James J. Heckman, Jora Stixrud, and Sergio Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482, 2006.
- Bengt Holmstrom and Paul Milgrom. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, 1991. URL <https://www.jstor.org/stable/764957>.
- Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):933–959, 2012.
- Sheena S. Iyengar and Emir Kamenica. Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7-8):530–539, 2010.
- C. Kirabo Jackson. Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. *The Economic Journal*, 120(549):1399–1429, 2010.
- C. Kirabo Jackson. Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *Review of Economics and Statistics*, 95(4):1096–1116, 2013.
- C. Kirabo Jackson. What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, 126(5):2072–2107, 2018.
- C. Kirabo Jackson. Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood? *Journal of Human Resources*, 2019.
- C. Kirabo Jackson, Shanette Porter, John Easton, Alyssa Blanchard, and Sebastián Kiguel. School Effects on Socio-emotional Development, School-Based Arrests, and Educational Attainment. *American Economic Review: Insights*, feb forthcoming.
- I T Jolliffe. *Principal Component Analysis, Second Edition*. Springer, 2002.

- Thomas Kane and Douglas Staiger. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. 2008.
- Tim Kautz, James J. Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. NBER Working Paper 20749, 2017.
- Lars J. Kirkeboen, Edwin Leuven, and Magne Mogstad. Field of Study, Earnings, and Self-Selection. *The Quarterly Journal of Economics*, 131(3):1057–1111, 2016.
- Jeffrey R Kling, Jeffrey B Liebman, and Lawrence F Katz. Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1):83–119, 2007.
- W. Bentley MacLeod and Miguel Urquiola. Reputation and School Competition. *American Economic Review*, 105(11):3471–3488, 2015.
- W. Bentley MacLeod and Miguel Urquiola. Is Education Consumption or Investment? Implications for the Effect of School Competition. NBER Working Paper 25117, Cambridge, MA, 2019.
- Richard K. Mansfield. Teacher Quality and Student Inequality. *Journal of Labor Economics*, 33(3):751–788, 2015.
- Isaac Mbiti, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani. Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania. *The Quarterly Journal of Economics*, 134(3):1627—1673, 2019.
- Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.
- Daniel McFadden. Conditional Logit Analysis of Qualitative Choice. In Ed. Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, 1973.
- Jack Mountjoy and Brent Hickman. The returns to college(s): Estimating value-added and match effects in higher education. *University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2020-08*, 1 2020. doi: 10.2139/SSRN.3537773. URL <https://papers.ssrn.com/abstract=3537773>.
- Hessel Oosterbeek, Nienke Ruijs, and Inge de Wolf. Using admission lotteries to estimate heterogeneous effects of elite schools. *SSRN Electronic Journal*, 4 2020. doi: 10.2139/SSRN.3580387. URL <https://papers.ssrn.com/abstract=3580387>.
- Parag A Pathak and Tayfun Sönmez. School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation. *American Economic Review*, 103(1): 80–106, 2013.
- Kate Place and Philip Gleason. Do Charter Middle Schools Improve Students’ College Outcomes? (Study Highlights). Technical report, Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2019.
- Cristian Pop-Eleches and Miguel Urquiola. Going to a Better School: Effects and Behavioral Responses. *American Economic Review*, 103(4):1289–1324, 2013.
- Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis (Econometric Society Monographs)*. 1992.
- Jesse M Rothstein. 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):1333–1350, 2006.
- Donald B Rubin. Formal Modes of Statistical Inference for Causal Effects. *Journal of Statistical Planning and Inference*, 25:279–292, 1990.
- Arthur Schram and Joep Sonnemans. How individuals choose health insurance: An experimental analysis. *European Economic Review*, 55(6):799–819, 2011.
- Barry Schwartz. *The paradox of choice: Why more is less*. HarperCollins Publishers, New York, NY, US, 2004. ISBN 0-06-000568-8 (Hardcover).
- Gary Solon, Steven J. Haider, and Jeffrey M. Wooldridge. What are we weighting for? *Journal of Human Resources*, 50:301–316, 3 2015. ISSN 0022-166X. doi: 10.3368/JHR.50.2.301. URL <http://jhr.uwpress.org/content/50/2/301><http://jhr.uwpress.org/content/50/2/301.abstract>.
- C. Spearman. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1):72, 1904.
- Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, 2009.

Appendix: NOT FOR PUBLICATION

Appendix A: School Placement Rules and Validity of the Regression Discontinuity Identification Strategy

The School Assignment Algorithm

School slots are assigned in rounds such that the most highly subscribed/ranked school fills its spots in the first round, then the next highly subscribed school fills its slots in the second round, and so on until all school slots are filled. This is done as follows: (1) the number of school slots at each school n_j is predetermined based on capacity constraints. (2) Each student is tentatively placed in the applicant pool for her first choice school and is ranked by SEA score. (3) The school at which the n_j^{th} ranked student has the highest SEA score is determined to be the most highly subscribed/ranked school and the top n_{j_1} students in the applicant pool for top-ranked school j_1 are assigned to school j_1 . The SEA score of the $n_{j_1}^{\text{th}}$ student is the cutoff score for school j_1 . (4) The top-ranked school slots and the assigned students are removed from the process, and the second choice becomes the new "first choice" for students who had the top-ranked school as their first choice but did not gain admission. (5) This process is repeated in round two to assign students to the second highest ranked school j_2 and determine the cutoff score for the second-ranked school, and this is repeated in subsequent rounds until all slots are filled. This assignment mechanism is a deferred acceptance algorithm (Gale and Shapley 1962) in which students have incentives to truthfully reveal their rankings among chosen schools.

However, there is an important exception to the school assignment algorithm-based rule. Specifically, Government assisted schools (which are privately managed public schools – akin to Charter schools in the US) can admit up to 20 percent of their incoming class at the principal's discretion. As such, the rule is used to admit at least 80 percent of the students at these schools, while the remaining students can be hand-picked by the principal before the next-highest ranked school fills any of its slots. For example, suppose the highest ranked school has 100 slots and is a Government assisted school. The top 80 applicants to that school will be admitted, while the principal can hand-pick up to 20 other students at her discretion. The remaining 20 students would be chosen based on for example family alumni connections, being relatives of teachers, religious affiliation, and so on. These hand-picked students may list the school as their top choice, but this need not be the case. Students receive one admission decision and are never made aware of other schools they would have been admitted to had they not been hand-picked. Only after all the spots (including both admitted students based on the algorithm and on the hand-picking) at the highest ranked school have been filled will the process be repeated for the remaining schools. As such, school admissions are based partly on the described deterministic function of student test scores and student choices and partly on the endogenous selection of students by school principals at Government assisted schools.

In addition, there are other circumstances by which the attended school would differ from the algorithm-based assigned school. First, students who do not score high enough to be assigned to a school on their choice list receive an administrative placement from the Ministry of Education (made to the administrative school zoned to the students' residential location). Finally, due to unforeseen circumstances some schools may have less capacity than expected or may close (this may happen due to flooding etc.). In such rare cases, the Ministry will place students to schools based on open slots in nearby schools, open slots in other schools in the choice list, and proximity.

Simulating the School Assignments Using the Algorithm-Based Rule

Because the assignment algorithm is known and we have the same data used by the Ministry of Education to tentatively assign students, we can identify the algorithm-based assignment cutoffs and, therefore, the algorithm-based school assignments (i.e. those that would have been the actual school allocations if Government assisted schools could not select any of their own students). This algorithm-based or tentative assignment removes the part of the actual admission process that may be driven by endogenous selection and leaves only the variation in the assignments that are known deterministic functions of students' test scores and school choices.

Following Jackson (2010) and Pop-Eleches and Urquiola (2013), we stack the data across all application pools for each year to each school (that is, we stack data for all the cutoffs into a single cutoff) into one single database. As such, we stack all application cutoffs and re-center the SEA scores for applicants to each school in each year around the algorithm-based assignment cutoff for that school-year.⁵⁵ Scoring above zero means scoring above the cutoff for a preferred school. Figure A1 shows the relationship between actually *attending* to one's preferred school as a function of one's incoming test score relative to the assignment cutoff for that school.⁵⁶ Consistent with our assignment cutoffs capturing real exogenous variation in actual school attendance, there is a sudden increase in the likelihood of attending a preferred school as one's score goes from below to above the assignment cutoff. Appendix Table A2 reports this first-stage estimated coefficient evidencing its high significance. This shows that there are meaningful differences in preferred school attendance associated with scoring above versus below an assignment cutoff that are not due to selection or hand-picking. Next, we provide direct supporting evidence on the exogeneity of the algorithm-based assignment cutoffs.

Testing the Exogeneity of the Assignment Cutoffs

The RD variation used in this paper is driven by the assignment cutoffs. As such, here we present evidence that this identification strategy is likely valid. One key diagnostic is to test for smoothness of density across the simulated cutoffs (McCrary 2008). As such, we formally test for any differential density across simulated cutoffs within each of our SEA cohorts by regressing the density of observations at each relative SEA score on an indicator for scoring above the cutoff along with smooth functions of the relative score.⁵⁷ We estimate this for each school cutoff in each year and report the average *t*-statistic associated with the null hypothesis of no differential density for each cohort along with the proportion of cutoffs that yield *p*-values smaller than 0.1. As one can see in Appendix Table A1, these tests reveal no statistically significant relationship between scoring above the cutoff and the density. Therefore, there is little evidence of gaming around the cutoffs regarding the density of observations at each test score.

⁵⁵Specifically, for each school we find all students who list that school as their top choice, re-center those students' SEA scores around the simulated cutoff for that school, and create a sample of applicants for each school. To mimic the sequential nature of the algorithm, we remove students assigned to their top choice schools, replace students' first choice with their second choice, and repeat this process with their second, third, fourth, fifth, and sixth choices. The applicant samples for all schools are then stacked so that every student has one observation for each school for which she/he was an applicant. We use four or six choices, as relevant per cohort limit. Only for SEA cohorts 2001-2006 students were allowed to list up to 6 school choices. Therefore, most of SEA cohorts in our data (1995-2000 and 2007-2012) could list up to 4 school choices.

⁵⁶We consider that one student attended school *j* if the student was enrolled in school *j* at the time of writing the CSEC examinations.

⁵⁷We implement these tests using the *rddensity* command in Stata.

The validity of the identification strategy also requires that there be no sorting of students around the cutoff (i.e. that latent outcomes are smooth through the cutoff). Given that students are unaware of the location of the cutoffs and are forced to make school choices before they take the SEA examinations, it is very unlikely that there is any sorting around the test score cutoffs. However, to provide further evidence that the variation employed (due to the cutoffs) is valid, we compute predicted outcomes (using the available baseline information) and test for whether scoring above the assignment cutoff is associated with any significant change in predicted outcomes.

Specifically, we first regress our outcomes on the number of SEA attempts (repeater status in 5th grade), the student’s sex, the student’s religion, selectivity of the student’s primary school (measured by the average SEA scores of each primary school-year), selectivity of the student’s secondary school choices (measured by the average SEA scores of the incoming class to each school choice-year), month of birth (to measure quarter of birth effects), age at SEA, and SEA cohorts fixed effects. These variables are relatively good predictors of the examination indexes such that, as shown in column 1 of [Appendix Table A2](#), they yield adjusted R-squares ranging from 0.27 to 0.31. However, the predictive power for the nonacademic binary outcomes is low.

We then take the fitted values from these prediction regressions as our predicted outcomes. If there was some gaming of the cutoff, one would likely see that scoring above the cutoff (conditional on smooth functions of the relative SEA score) should be associated with better “predicted” scores. However, with no gaming there should be no relationship between scoring above the cutoff and one’s predicted outcomes. To test for this, we estimate the following model using our stacked database:

$$Y_{ijt}^p = \pi \cdot Above_{i\tau t} + f(SEA_{it}) + F_{\tau t} + \varepsilon_{ijt} \quad (17)$$

where Y_{ijt}^p is the predicted outcome for individual i who attended school j at time t . $Above_{i\tau t}$ is an indicator for scoring above the algorithm-based assignment cutoff for school τ . Among those who comply with the cutoff, $j=\tau$. $f(SEA_{it})$ is a 5th order polynomial of the incoming SEA score net of the cutoff score for preferred school τ . $F_{\tau t}$ is a cutoff fixed effect for applicants to school τ in year t . The inclusion of cutoff fixed effects ensures that all comparisons are among students who applied to the same school in the same year. Because the same individual can enter the data for multiple cutoffs, the estimated standard errors are clustered at the individual level.

Consistent with no gaming, column 2 of [Appendix Table A2](#), shows that there is no relationship between scoring above the cutoff and one’s predicted outcomes. The estimated coefficients, $\hat{\pi}$, are small in magnitude and statistically indistinguishable from zero – indicating no gaming across the assignment cutoffs. As an additional check on this model, we estimated model (17) for different bandwidths around the cutoff. [Figure A2](#) presents these results visually. As one can see for any choice of bandwidth, there are no effects of scoring above the cutoff on predicted outcomes. Taken together, the patterns suggest that the variation due to the algorithm-based assignment cutoffs is likely exogenous and, therefore, valid to identify causal school impacts.

Finally, [Appendix Table A2](#) also reports the estimated RD effects on the actual outcomes (columns 4 -7) showing that reduced-form effects of scoring above the school assignment cutoff are associated, on average, with significant improvements in students’ examination indexes and that these estimates are not sensitive to the inclusion of baseline sociodemographic controls in the model.

Reconciling Null Pooled Cutoff Impacts with Other Results

One may wonder how scoring above a cut-off for a preferred school on average could be unrelated to outcomes while there are real differences in value-added across schools. This could occur if (a) school choices are unrelated to value-added, (b) school preferences are heterogeneous, or (c) there is little variability in value-added among choices. We discuss each below.

- (a) Consider the hypothetical scenario where school choices are entirely random. In this hypothetical scenario, preferred schools are not higher value-added on average, and there would be no pooled cut-off effect even if there are real differences in value-added across schools. This hypothetical scenario highlights that there could be no aggregate cut-off effects if preferences for school (among those at the cut-offs) are not strongly related to value-added in all dimensions. This example is not meant to explain why the results differ between Table 2 and Table A2, but rather to underscore the fact that a lack of an aggregate cut-off effect does not imply a lack of school effects.
- (b) A more plausible explanation for small cut-off effects despite there being real differences in value-added is that preferences for schools are heterogeneous. That is, some people may rank high value-added schools more highly, while others do not so that on average the pooled threshold effect is small. Indeed, we document clear differences in the choice behaviors of parents of children at the top versus the bottom of the incoming test score distribution so that this is not unreasonable.
- (c) Another possibility is that many parents may choose schools that are very similar in value-added along those dimensions. For example, a parent who wishes to have their child attend a school that reduces crime may list four schools that are all high value-added in that dimension. In such a case, there are real differences in school value-added, choices are heavily influenced by school value-added, but there would be no effect on value-added through an admission threshold. This scenario may hold (or approximately hold) for many parents in the data such that many individual cut-offs end up having very low signal-to-noise ratios.

A related question (in light of scenario (a) above) is whether the lack of an aggregate cut-off effect implies that our choice models must be wrong (i.e., that parents do choose schools with impacts on non-academic outcomes conditional on proximity, admission probability, peer characteristics, and average outcomes). We argue that it does not for three reasons.

1. As we discuss above, the pooled average results may conceal considerable heterogeneity and could be explained by the heterogeneity in preferences for schools that are suggested by our choice patterns (scenario b above).
2. Also, given scenario (c) above, if parents have strong preferences for an attribute, the schools listed may all have similar levels of value-added for that attribute. In this scenario, that attribute will strongly predict being listed as a top choice school, but there may be little difference in value-added among the choices made (and thus no cut-off effect in the RD model).

3. It is important to point out that our choice models show relationships that are conditional on peer quality, proximity, school-level averages and include school impacts in multiple dimensions. As such, it is possible that the preferred school is not a higher value-added school unconditionally (which is what Table 2 will speak to), but is only higher value-added conditional on all the other predictors of choices. To give a hypothetical example, suppose John values proximity and arrests value-added. In the choice models with both included, one would find that John is more likely to rank schools with high proximity and high value-added. However, if for some reason, John lives far away from the high value-added schools and has a heavy weight on proximity, there may be little difference in the arrests value-added of John's choices (and he may even list a lower value-added school above a higher one). This example is merely meant to illustrate that because the choice models represent conditional relationships, the patterns we document are not directly relatable to the pooled cut-off effects.

Table A1. Testing for differential density around the school assignment cutoff

SEA Cohort	<i>p-value</i> < .1	<i>mean</i> <i>T-stat</i>	SEA Cohort	<i>p-value</i> < .1	<i>mean</i> <i>T-stat</i>
1995	9.46%	-0.04	2004	12.77%	-0.19
1996	10.81%	-0.02	2005	7.69%	0.03
1997	16.22%	-0.16	2006	11.46%	-0.02
1998	10.53%	-0.09	2007	11.70%	-0.02
1999	17.11%	0.11	2008	9.89%	0.22
2000	11.11%	0.06	2009	5.43%	-0.04
2001	11.58%	-0.05	2010	11.39%	-0.08
2002	13.68%	0.18	2011	8.86%	-0.01
2003	15.96%	0.21	2012	15.79%	0.21

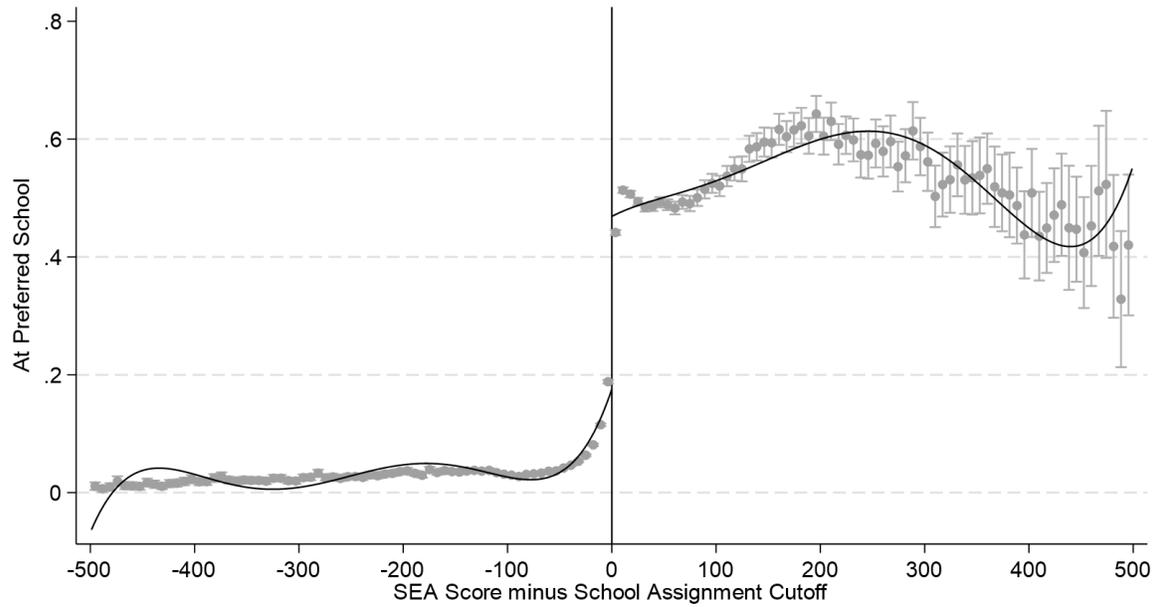
Notes: This table shows the percentage of school assignment cutoffs for which the *p-value* corresponding to the test of differential density across the cutoff is less than 10% in each SEA cohort. For all cohorts together it is 11.69%. The table also reports the average of *T-Statistics* of differential density tests of all school assignment cutoffs for each SEA cohort included in the study.

Table A2. First Stage and Reduced-Form Effects

	Predicted Outcomes			Actual Outcomes			
	Prediction R2	Effect	<i>p-value</i>	Effect	<i>p-value</i>	Effect	<i>p-value</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First Stage:							
Attended preferred school				0.200	<0.001	0.200	<0.001
Reduced-Form Effects:							
High-Stakes Index	0.31	-0.003	0.173	0.062	<0.001	0.062	<0.001
Low-Stakes Index	0.27	0.000	0.968	0.048	<0.001	0.050	<0.001
No Dropout by 14	0.10	-0.001	0.095	-0.004	0.08	-0.004	0.07
No live birth by 19	0.03	0.000	0.578	-0.001	0.86	-0.001	0.79
Not arrested by 18	0.04	0.000	0.634	0.000	0.83	0.000	0.86
Formally employed 27+	0.02	-0.001	0.173	-0.004	0.38	-0.004	0.36
Cutoff fixed effects		Yes		Yes		Yes	
Sociodemographics		No		No		Yes	

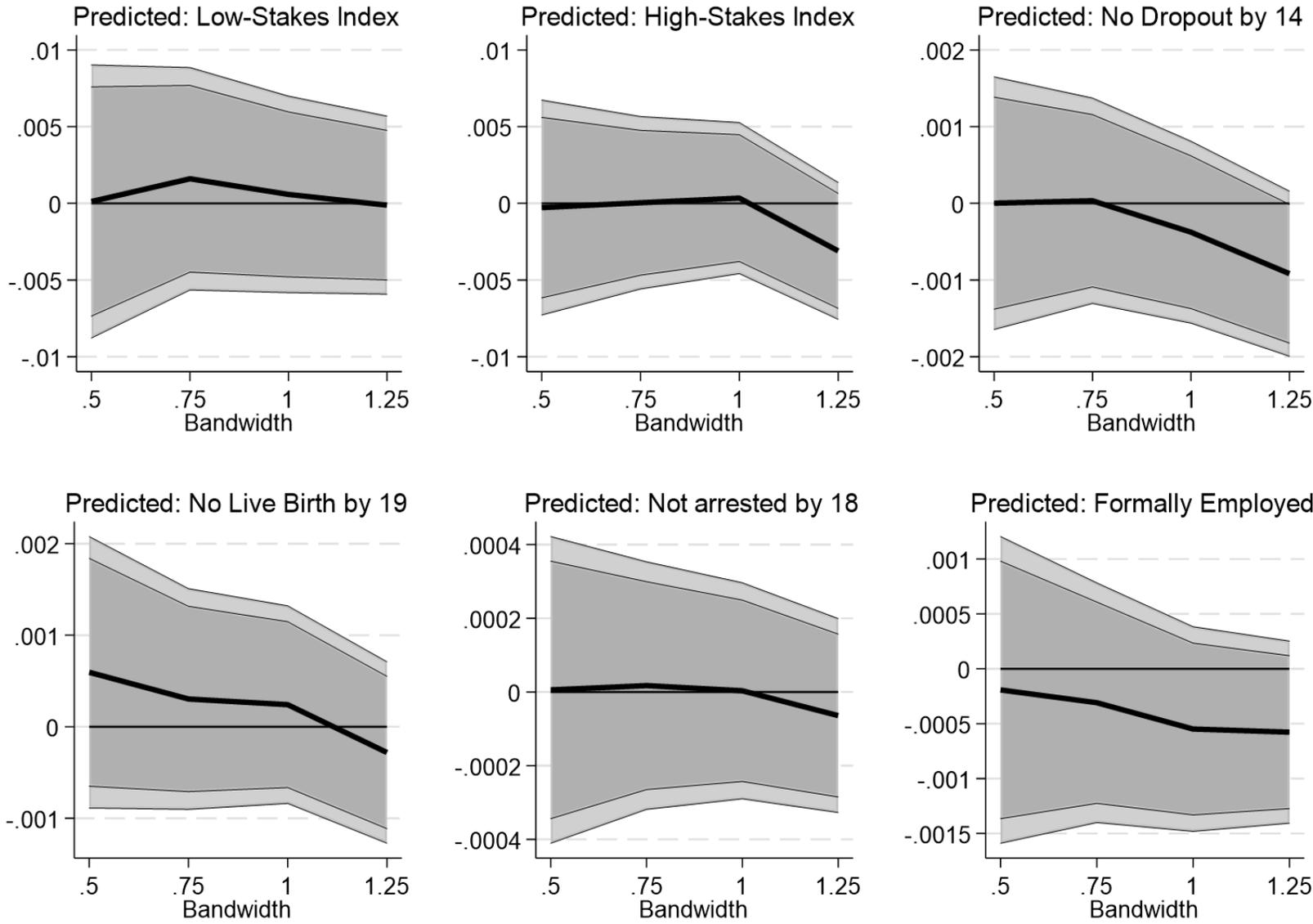
Notes: This table reports estimated coefficients on 'Above' from equation (17). Models were estimated using observations within a bandwidth of +/-1.25 standard deviations from the school assignment cutoff. Sociodemographics include sex, primary school district fixed effects, and religion fixed effects. Estimated standard errors are clustered at the individual level in all regressions. *P-values* for the null of $\pi=0$ shown next to the estimated coefficients.

Figure A1. Discontinuity in Preferred School Attendance Through Assignment Cutoffs



Notes: The Y-axis represents the likelihood of preferred school attendance (i.e. the school where the student was enrolled at the time of taking the CSEC examinations). The X-axis is the SEA score relative to the deferred acceptance rule-based assignment cutoff. The circles are means corresponding to 7-point bins of the relative score. The solid lines are the fitted school attendance rates generated by fitting a fifth degree polynomial of the relative score fully interacted with an indicator for scoring above the school assignment cutoff. The gray vertical bars depict the 90 percent confidence intervals for each bin average.

Figure A2. Reduced-form effects on predicted outcomes by bandwidth



Notes: This figure reports estimated coefficients on 'Above' resulting from equation (17). The estimated coefficients are reported for each bandwidth between $\pm 0.5sd$ and $\pm 1.25sd$ from the school assignment cutoff. The 90 (95) percent confidence intervals of the estimated coefficients are presented in dark (light) gray.

Appendix B: Appendix Tables and Figures

Table B1. Minimum and maximum cutoffs by district

District	Cutoff by SEA percentile	
	Minimum	Maximum
	(1)	(2)
Caroni	4	97
North Eastern	1	82
Port of Spain	7	97
South Eastern	1	91
St. George East	1	98
St. Patrick	1	92
Tobago	3	77
Victoria	1	98

Notes: We show the minimum and maximum cutoffs as SEA percentiles for each district across all years.

Table B2: Weights Used to Compute Indexes

High-Stakes Index	Weight
Number of CSEC subjects passed	0.202
CSEC tertiary qualification	0.192
CSEC tertiary qualification attempt	0.140
CAPE scholarship	0.068
CAPE scholarship attempt	0.213
Number of CAPE units passed	0.219
CAPE Associate's degree	0.213

Low-Stakes Index	Weight
NCSE Total Academic	0.546
NCSE Total Non academic	0.546

Notes: Indexes are computed from a separate factor analysis (using the principal-component factor method) applied to the individual outcomes that integrate each index. The weights for individual outcomes within the indexes are determined by predicting the first underlying principal-component applied separately to each group of outcomes that integrate each index. The computed indexes are standardized to have zero mean and unit variance. CSEC tertiary qualification is obtained when passing 5 subjects including English language and mathematics. “CSEC tertiary qualification attempt” denotes that the student took 5 subjects including English language and mathematics. CAPE scholarship is awarded when passing eight CAPE units (including Caribbean and Communication studies) with the maximum possible grade. “CAPE scholarship attempt” denotes that the student took eight CAPE units (including Caribbean and Communication studies). CAPE associate's degree is awarded when passing seven CAPE units (including Caribbean and Communication studies). NCSE academic subjects include mathematics, English, Spanish, sciences, and social studies. NCSE non academic subjects include arts, physical education, and technical studies.

Table B3: Two sample test for average and weighted 2SLS

	Assuming Correlation = 0			Assuming Correlation = 1		
	Avg - 25th	Avg - 50th	Avg - 75th	Avg - 25th	Avg - 50th	Avg - 75th
	(1)	(2)	(3)	(4)	(5)	(6)
All Outcomes	0.00	0.00	0.00	0.00	0.00	4.63
High-Stakes Index	0.00	0.00	0.00	0.00	0.00	0.00
Low-Stakes Index	0.00	0.00	0.00	0.00	0.00	2.02
No Dropout by 14	0.00	0.00	0.00	0.00	0.00	8.25
Not arrested by 18	0.00	0.00	0.00	0.00	0.00	16.49
No live birth by 19	0.00	0.00	0.00	0.00	0.00	0.00
Formally employed 27+	0.00	0.00	0.00	0.00	0.00	0.00

Notes: We report the percentage of schools where the difference in the average and weighted value-added is significant at the 95% level. We adjust the difference dividing it by its standard error, which is estimated for the following two scenarios: when the correlation between the estimated value-added is assumed to be 0 and when it is assumed to be 1. This will affect the variance of the difference through the covariance term. We then report the percentage of cases where the adjusted difference is greater than 1.96.

Table B4: Testing for Differential Match Effects

	Mean	p-value	Median	p-value
	(1)	(2)	(3)	(4)
High-Stakes Index	0.702	0.323	0.015	0.179
Low-Stakes Index	0.2	0.146	-0.016	0.486
No Dropout by 14	0.051	0.502	0.006	0.735
Not arrested by 18	-2.939	0.304	0.003	0.558
No live birth by 19	-2.199	0.1	-0.036	0.185
Formally employed 27+	0.08	0.295	0.015	0.443

Notes: We show results for a mean and median test for the interaction effect, $\theta_{k,m}^{Pref}$, from (7). Neither is rejected for being equal to zero at the 10% level. Hence, we find no evidence of *differential effect* of being assigned to school m relative to k for those who chose k before m relative to those who listed m before k.

Table B5: Standard Deviation of Persistent School Impacts and Maximum Likelihood Correlations Between 2SLS School Impacts

Outcome	School Level (σ_{error})				School Level (correlations)			
	Size of Impact	Size of impacts for Students at the			Correlation	Correlations with High-Stakes at the		
		25th %ile of the achievement distribution	median of the achievement distribution	75th %ile of the achievement distribution		25th %ile of the achievement distribution	median of the achievement distribution	75th %ile of the achievement distribution
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Standardized Outcomes								
<i>High-Stakes Index</i>								
All Schools	0.441 [0.397 , 0.483]	0.516 [0.464 , 0.562]	0.495 [0.445 , 0.534]	0.444 [0.400 , 0.481]	1.000	1.000	1.000	1.000
Dropping Outliers	0.441 [0.397 , 0.483]	0.516 [0.464 , 0.562]	0.495 [0.445 , 0.534]	0.444 [0.400 , 0.481]	1.000	1.000	1.000	1.000
<i>Low-Stakes Index</i>								
All Schools	0.493 [0.452 , 0.530]	0.864 [0.715 , 0.962]	0.814 [0.759 , 0.861]	0.441 [0.389 , 0.481]	0.091 [-0.045 , 0.219]	0.102 [-0.028 , 0.253]	0.160 [0.037 , 0.257]	0.000 [-0.136 , 0.141]
Dropping Outliers	0.473 [0.433 , 0.508]	0.764 [0.683 , 0.826]	0.814 [0.760 , 0.859]	0.410 [0.365 , 0.448]	0.100 [-0.042 , 0.224]	0.173 [0.047 , 0.289]	0.173 [0.041 , 0.285]	0.037 [-0.110 , 0.167]
<i>Low-Stakes Index (Academic)</i>								
All Schools	0.390 [0.360 , 0.414]	0.699 [0.637 , 0.754]	0.608 [0.566 , 0.646]	0.324 [0.283 , 0.357]	0.196 [0.076 , 0.312]	0.184 [0.047 , 0.292]	0.147 [0.020 , 0.260]	0.207 [0.086 , 0.337]
Dropping Outliers	0.390 [0.360 , 0.414]	0.699 [0.637 , 0.754]	0.608 [0.566 , 0.646]	0.324 [0.283 , 0.357]	0.196 [0.076 , 0.312]	0.184 [0.047 , 0.292]	0.147 [0.020 , 0.260]	0.207 [0.086 , 0.337]
<i>Low-Stakes Index (Non Academic)</i>								
All Schools	0.618 [0.560 , 0.664]	0.950 [0.778 , 1.055]	0.971 [0.898 , 1.032]	0.600 [0.538 , 0.653]	0.011 [-0.141 , 0.145]	0.068 [-0.067 , 0.214]	0.151 [0.035 , 0.275]	-0.098 [-0.248 , 0.037]
Dropping Outliers	0.607 [0.549 , 0.657]	0.831 [0.743 , 0.903]	0.970 [0.898 , 1.036]	0.597 [0.542 , 0.650]	0.024 [-0.129 , 0.159]	0.132 [-0.014 , 0.262]	0.166 [0.046 , 0.287]	-0.087 [-0.229 , 0.047]

Notes: All estimates shown were computed by bootstrap with 1,000 repetitions of the maximum likelihood approach described in the text. We report the median as the point estimate, as well as the 5th and 95th percentiles for the confidence intervals. Columns (1) - (4) report estimated standard deviations of the persistent school impacts for each outcome; where school impacts have been estimated without weights (column 1), and with weights centered at the 25th, 50th and 75th percentile of the incoming achievement distribution (columns 2 to 4 respectively). Column (5) - (8) report estimated correlations of the persistent school impacts on the high-stakes index with the persistent school impacts on other outcomes; where school impacts have been estimated without weights (column 5), and with weights centered at the 25th, 50th and 75th percentile of the incoming achievement distribution (columns 6 to 8 respectively). We do this using $weight_i = (1 + \frac{(X - pct_i)^2}{100})^{-1}$, where $X = 25, 50, 75$ and pct_i is the student's percentile in the achievement distribution. For the first row of each outcome, we removed schools with outlier estimated impacts (i.e. beyond 4σ of the median school). Estimates reported in the second row of each outcome were obtained without removing outliers.

Table B5 (continued): Standard Deviation of Persistent School Impacts and Maximum Likelihood Correlations Between 2SLS School Impacts

Outcome	School Level ($\sigma_{\theta error}$)				School Level (correlations)			
	Size of Impact	Size of impacts for Students at the			Correlation	Correlations with High-Stakes at the		
		25th %ile of the achievement distribution	median of the achievement distribution	75th %ile of the achievement distribution		25th %ile of the achievement distribution	median of the achievement distribution	75th %ile of the achievement distribution
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Binary Outcomes								
<i>No Dropout by 14</i>								
All Schools	0.148 [0.091 , 0.162]	0.197 [0.175 , 0.215]	0.135 [0.089 , 0.146]	0.170 [0.074 , 0.183]	0.117 [-0.009 , 0.266]	0.002 [-0.166 , 0.135]	0.102 [-0.065 , 0.354]	0.238 [0.006 , 0.487]
Dropping Outliers	0.090 [0.077 , 0.100]	0.190 [0.171 , 0.206]	0.090 [0.081 , 0.098]	0.070 [0.058 , 0.080]	0.121 [0.019 , 0.228]	-0.012 [-0.142 , 0.117]	0.238 [0.121 , 0.345]	0.117 [-0.044 , 0.290]
<i>Not arrested by 18</i>								
All Schools	0.068 [0.041 , 0.077]	0.072 [0.056 , 0.083]	0.134 [0.051 , 0.150]	0.075 [0.033 , 0.083]	0.084 [0.004 , 0.262]	-0.099 [-0.235 , 0.182]	-0.006 [-0.112 , 0.199]	-0.085 [-0.220 , 0.257]
Dropping Outliers	0.037 [0.032 , 0.041]	0.053 [0.048 , 0.059]	0.045 [0.039 , 0.050]	0.020 [0.017 , 0.023]	0.282 [0.164 , 0.419]	0.055 [-0.059 , 0.205]	0.097 [-0.022 , 0.226]	0.407 [0.210 , 0.619]
<i>No live birth by 19</i>								
All Schools	0.367 [0.295 , 0.414]	0.287 [0.254 , 0.311]	0.324 [0.200 , 0.369]	0.450 [0.339 , 0.527]	-0.022 [-0.115 , 0.092]	0.030 [-0.110 , 0.151]	-0.115 [-0.230 , -0.004]	-0.079 [-0.260 , 0.099]
Dropping Outliers	0.173 [0.147 , 0.193]	0.287 [0.254 , 0.313]	0.202 [0.165 , 0.232]	0.131 [0.107 , 0.150]	-0.036 [-0.171 , 0.087]	0.025 [-0.124 , 0.161]	-0.064 [-0.188 , 0.071]	0.174 [0.021 , 0.311]
<i>Formally employed 27+</i>								
All Schools	0.160 [0.078 , 0.191]	0.132 [0.086 , 0.149]	0.453 [0.093 , 0.484]	0.306 [0.115 , 0.349]	0.076 [-0.014 , 0.233]	0.193 [0.091 , 0.298]	-0.429 [-0.493 , 0.184]	-0.049 [-0.245 , 0.308]
Dropping Outliers	0.070 [0.058 , 0.079]	0.083 [0.065 , 0.098]	0.086 [0.072 , 0.097]	0.090 [0.076 , 0.103]	0.152 [0.025 , 0.294]	0.165 [0.046 , 0.297]	0.106 [-0.043 , 0.237]	0.014 [-0.148 , 0.157]

Notes: All estimates shown were computed by bootstrap with 1,000 repetitions of the maximum likelihood approach described in the text. We report the median as the point estimate, as well as the 5th and 95th percentiles for the confidence intervals. Columns (1) - (4) report estimated standard deviations of the persistent school impacts for each outcome; where school impacts have been estimated without weights (column 1), and with weights centered at the 25th, 50th and 75th percentile of the incoming achievement distribution (columns 2 to 4 respectively). Column (5) - (8) report estimated correlations of the persistent school impacts on the high-stakes index with the persistent school impacts on other outcomes; where school impacts have been estimated without weights (column 5), and with weights centered at the 25th, 50th and 75th percentile of the incoming achievement distribution (columns 6 to 8 respectively). We do this using $weight_i = (1 + \frac{(X - pct_i)^2}{100})^{-1}$, where $X = 25, 50, 75$ and pct_i is the student's percentile in the achievement distribution. For the first row of each outcome, we removed schools with outlier estimated impacts (i.e. beyond 4σ of the median school). Estimates reported in the second row of each outcome were obtained without removing outliers.

Table B6. Standard Deviation of Persistent School Impacts and even-odd year Correlations Between 2SLS School Impacts

Outcome	School Level ($\sigma_{\theta_j^{TOR}}$)	School Level correlations with high-stakes	
	Size of Impact	Average	75th %ile of the achievement distribution
	(1)	(2)	(3)
Standardized outcomes			
High-Stakes Index	0.432 [0.373 , 0.488]	1.000	1.000
Low-Stakes Index	0.454 [0.406 , 0.505]	0.105 [-0.066 , 0.272]	0.073 [-0.146 , 0.271]
Binary outcomes			
No Dropout by 14	0.084 [0.069 , 0.099]	0.088 [-0.050 , 0.228]	0.183 [-0.103 , 0.413]
Not arrested by 18	0.034 [0.028 , 0.040]	0.272 [0.112 , 0.418]	0.353 [0.117 , 0.604]
No live birth by 19	0.149 [0.117 , 0.178]	-0.001 [-0.180 , 0.163]	0.172 [-0.051 , 0.447]
Formally employed 27+	0.053 [0.035 , 0.069]	0.291 [0.055 , 0.536]	-0.013 [-0.246 , 0.215]

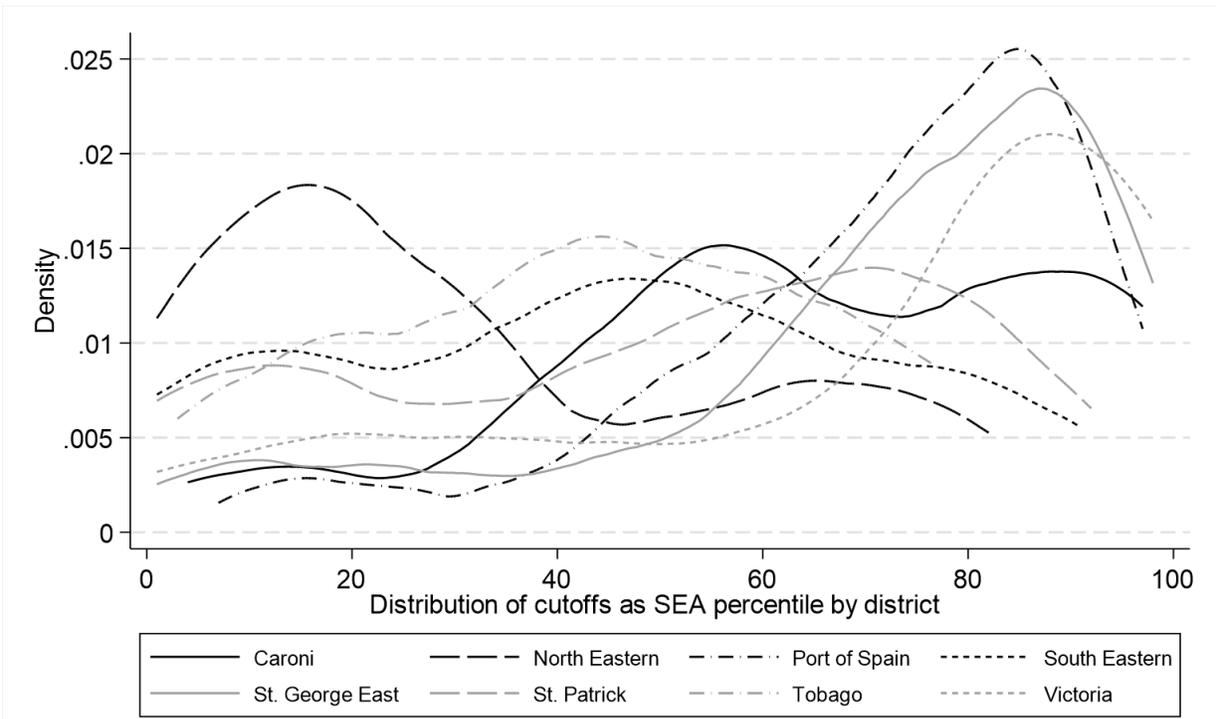
Notes: The table reports standard deviations and correlations of persistent school effects across different outcomes. Correlations were computed comparing even and odd years as described in the text. To disentangle this raw correlation, one must divide by the square root of the product of the reliability ratios for each measure. Correlations in the second column were computed with estimated school impacts for the average student and those in the third column were computed using school impacts with weights centered around the 75th percentiles of the achievement distribution. We do this using $weight_i = (1 + \frac{(75-pct_i)^2}{100})^{-1}$, where pct_i is the student's percentile in the achievement distribution. We removed schools with outlier estimated impacts (i.e. beyond 4σ of the median school). We bootstrap the even-odd correlations approach for 1,000 repetitions. Results shown correspond to the median (for the point estimates) and the 5th and 95th percentiles for the confidence intervals.

Table B7. Maximum Likelihood Correlations Using Different Weights for 2SLS estimates

	Correlation between average school impacts and school impacts weighted around the		
	25th %ile of the achievement distribution	median of the achievement distribution	75th %ile of the achievement distribution
	(1)	(2)	(3)
High-Stakes Index	0.99 [0.96 , 1.00]	0.99 [0.98 , 1.00]	0.97 [0.95 , 0.99]
Low-Stakes Index	0.90 [0.86 , 0.94]	0.91 [0.88 , 0.94]	0.91 [0.86 , 0.94]
No Dropout by 14	0.95 [0.86 , 1.00]	0.86 [0.80 , 0.93]	0.52 [0.34 , 0.65]
Not arrested by 18	0.93 [0.86 , 0.99]	0.73 [0.65 , 0.80]	0.74 [0.60 , 0.85]
No live birth by 19	0.92 [0.87 , 0.96]	0.89 [0.77 , 0.99]	0.93 [0.82 , 1.00]
Formally employed 27+	0.99 [0.85 , 1.00]	1.00 [0.94 , 1.00]	0.92 [0.83 , 1.00]

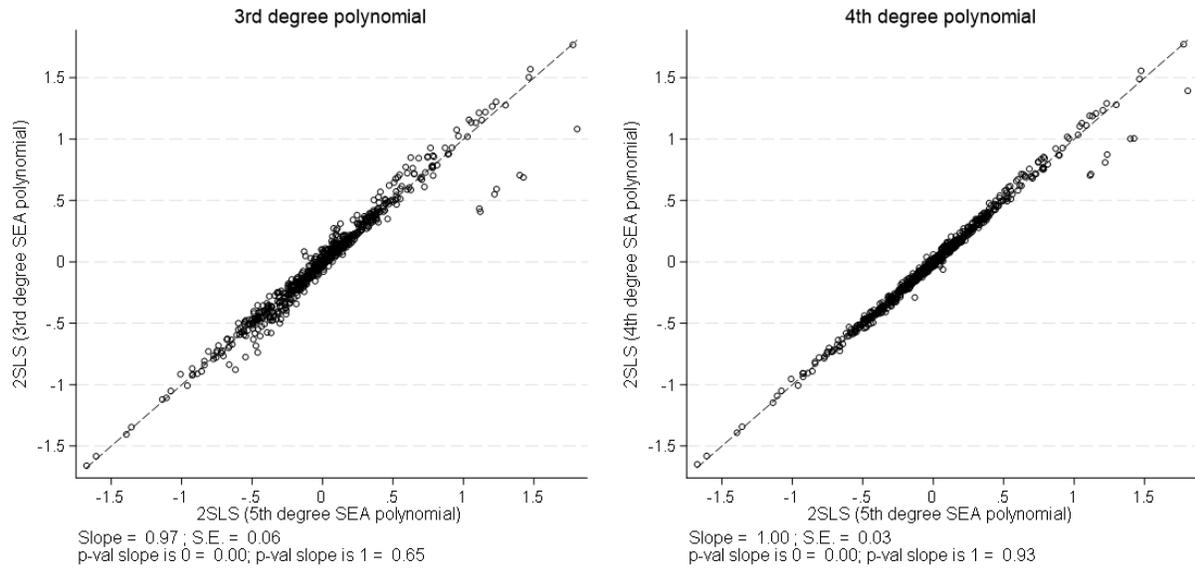
Notes: The table reports correlations of persistent school effects for each outcome across the different weighted estimates. Correlations were computed using the maximum likelihood approach described in the text. We show correlations between estimated school effects for the average student and those weighted around the 25th percentile, the median, and the 75th percentile of the students' SEA score distribution. Weighted school impacts are estimated using $weight_i = (1 + \frac{(X - pct_i)^2}{100})^{-1}$, where $X = 25, 50, 75$ and pct_i is the student's percentile in the achievement distribution. We removed schools with outlier estimated impacts (i.e. beyond 4σ of the median school). We bootstrap the maximum likelihood approach for 1,000 repetitions. Results shown correspond to the median (for the point estimates) and the 5th and 95th percentiles for the confidence intervals.

Figure B1. Distribution of cutoffs by district



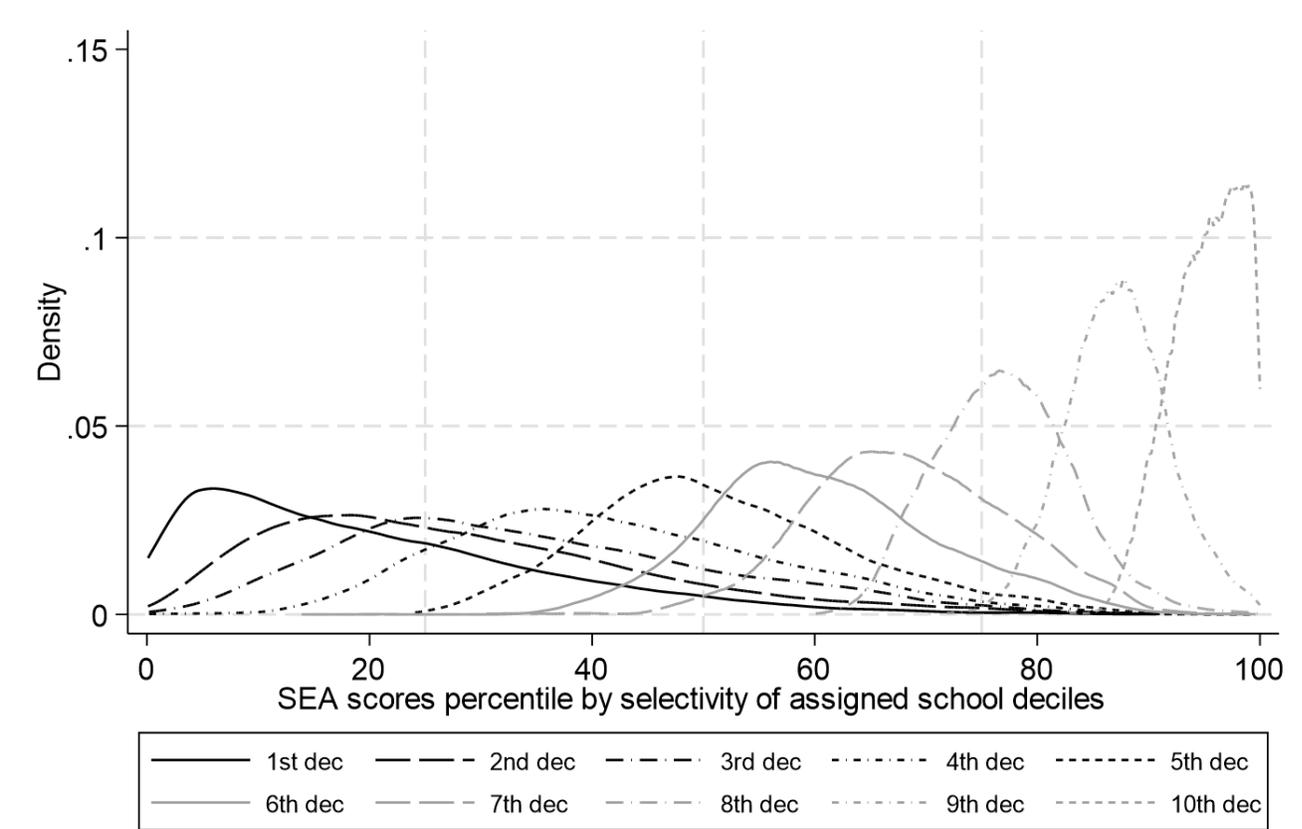
Notes: We show the distribution of cutoffs for all school-years by district. Cutoffs are shown as the SEA percentile they correspond to in each year.

Figure B2. Estimated School Impacts with alternative SEA polynomials



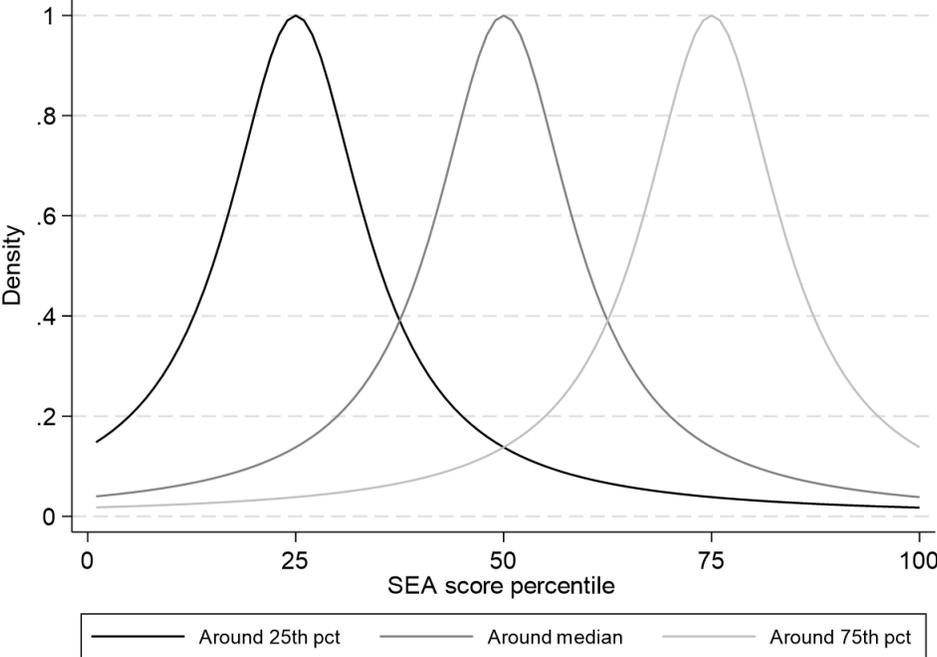
Notes: The left panel displays estimated 2SLS school impacts with a 3rd degree polynomial of the SEA score (Y-axis) against the preferred estimated school impacts with a 5th degree polynomial of the SEA score (X-axis). The right panel displays estimated 2SLS school impacts with a 4th degree polynomial of the SEA score (Y-axis) against the preferred estimated school impacts with a 5th degree polynomial of the SEA score (X-axis). Estimated slopes and tests for whether they differ from 0 and 1 are shown below each panel.

Figure B3. Common support of incoming scores by selectivity of assigned school (10% groups)



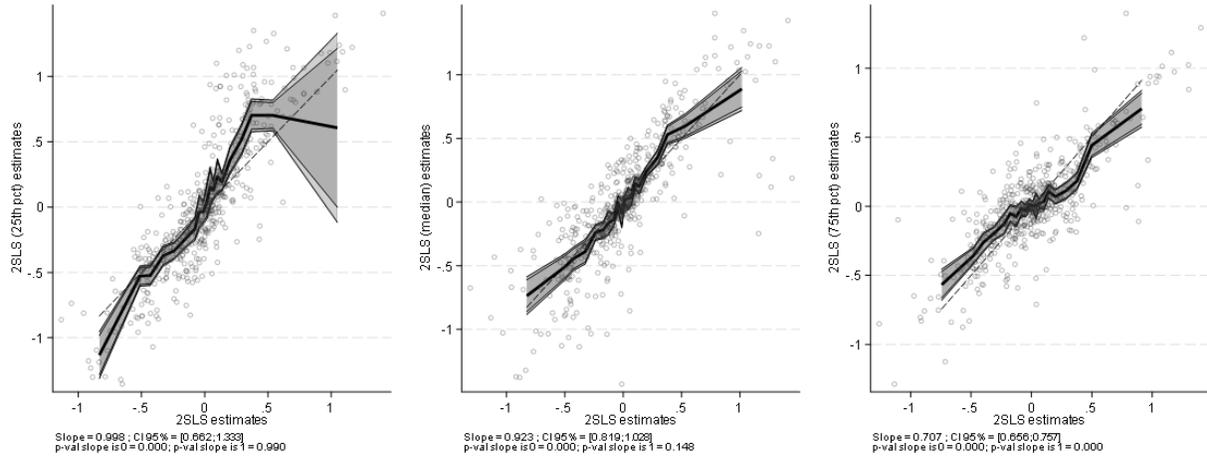
Notes: The X-axis represents the SEA score percentile. The Y-axis represents the density of students for each SEA score percentile. We plot one curve for different levels of school selectivity, going from the least selective on the left to the most selective schools on the right. As expected, as the selectivity of the school increases, the distribution of students by SEA percentiles shifts to the right. However, we can see there is still overlap of students between less and more selective schools.

Figure B4. Distribution of weights centered around 25th, median and 75th percentile



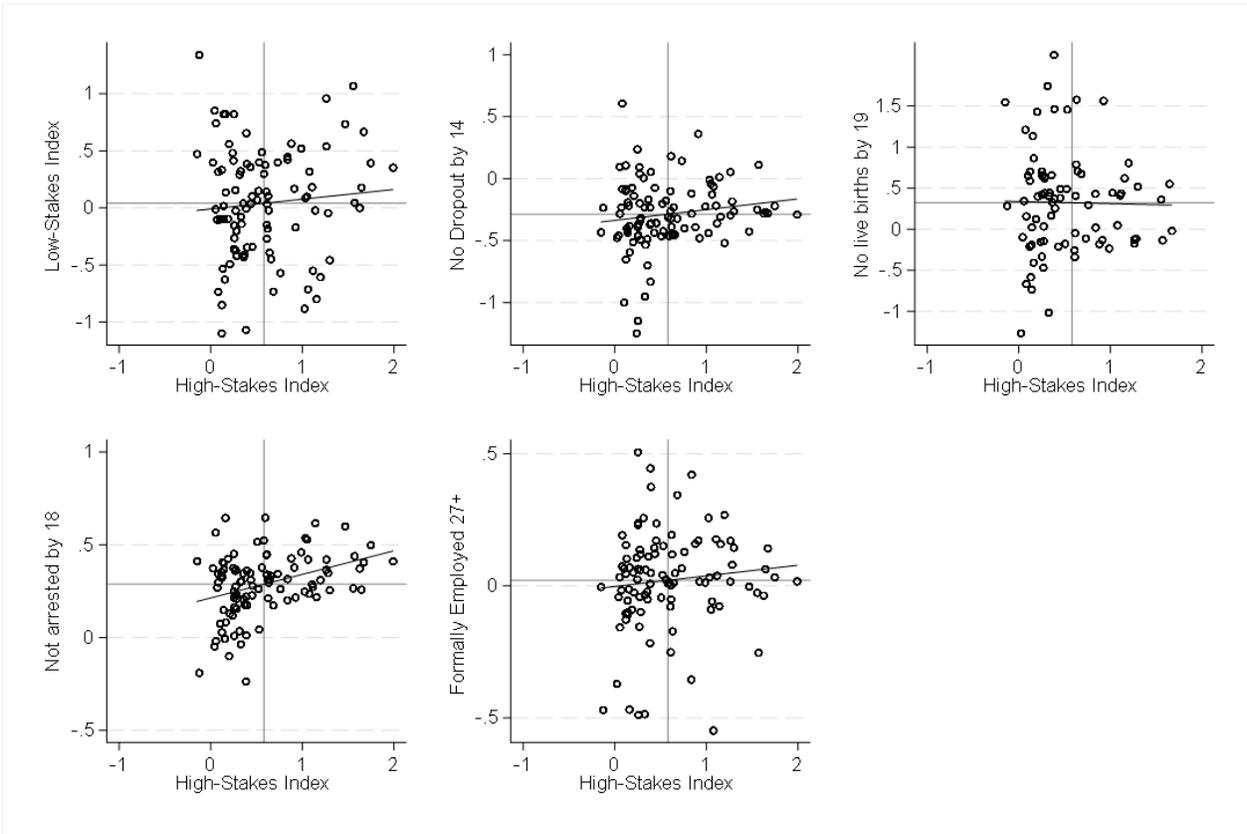
Notes: The X-axis represents the SEA score percentile. The Y-axis represents the density of students for each SEA score percentile. Each of the three weights is highly concentrated around the respective percentile (25th, 50th, 75th) but includes the whole range of students.

Figure B5. Predicted weighted 2SLS effects from average 2SLS effects



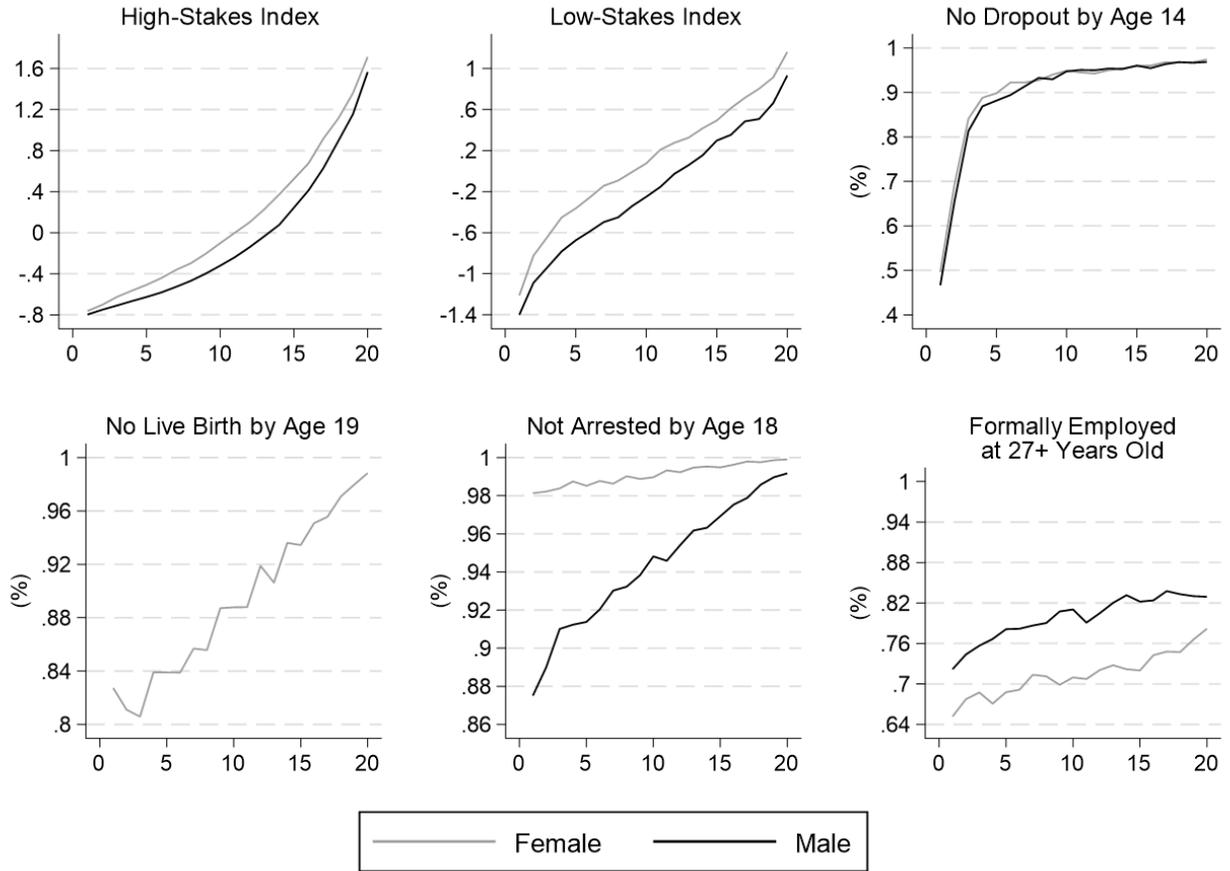
Notes: We plot weighted 2SLS estimated school impacts (around the 25th percentile, median and 75th percentile) in the Y-axis against the average 2SLS estimated school impacts (X-axis) in each panel. Prediction intervals have been computed by grouping the school estimates in 20 bins across the X-axis.

Figure B6. Scatter plots of the raw school effects



Notes: The X-axis plots the High-Stakes value-added for each school. The Y-axis represents the value-added for each school for the other outcomes.

Figure B7. Outcomes by SEA score ventile



Notes: The X-axis represents the SEA score ventile. The Y-axis represents each outcome for students in standardized units for high-stakes and low-stakes indexes and in percentages for the other outcomes. Arrests are much higher in males and in the lower ventiles, but it still occurs in the top ventiles of the SEA score.

Appendix C: Selection of excluded schools

Our analysis focuses on all public secondary schools in Trinidad and Tobago between the years 1995 and 2012. Therefore, all private schools, temporary schools, and non-traditional secondary schools (such as junior life centers that focus on life skills rather than the national academic curriculum) are excluded. Our analysis therefore focuses on 134 public secondary schools (which account for over 95 percent of student enrolment).

In our 2SLS models, we estimate one first stage regression for each of these 134 schools, where the school attendance indicator is the dependent variable; while each of the 134 schools algorithm-based *assignment* indicators plus all controls listed in equation (5) enter as regressors. For each school, we assess the point estimates on the own-school algorithm-based assignment indicator from each first stage regression. If the the own-school assignment indicator resulted is less than a 15 percentage-point increase in the likelihood of attendance, the school was identified as having a weak fist stage and we, therefore, excluded both the assignment and attendance indicators when estimating individual school's causal impacts. In addition, for those outcomes that are only observed for more recent cohorts (such as low stakes, dropout, and pregnancy which are only observed for cohorts after 2004 compared to 1995 for crime and formal employment, and 1999 for high stakes exams), we drop the next lowest compliance school in Tobago.

This threshold of 15 percentage points was chosen because lower thresholds resulted in the inclusion of schools that triggered weak identification for all schools. Using higher thresholds yielded very similar results. Therefore, we used the lowest threshold for which the full 2SLS model was well identified. Because some outcomes are not observed for all SEA cohorts, the first stages are somewhat different. As such, in order to avoid weak identification (only for those outcomes observed with recent cohorts only) an additional low-compliance school in Tobago needed to be dropped to avoid weak identification.

We do not exclude any individual observations when estimating the final 2SLS model. Therefore, the low-compliance schools for which the assignment and attendance indicators were excluded serve as the omitted category. Note that because not all schools were operating across during the entire analysis period, and because not all outcomes are observed for all the SEA cohorts; the total number of schools for which value-added estimates are observable varies across outcomes.

After estimating the school effects for each outcome, the main analysis removes any remaining outlier estimates (those more than 4σ away from the median estimate for that outcome). [Appendix Table C1](#) shows the total number of schools for which we can observe each outcome (column 1), the total number of schools for which we estimate value-added for each outcome (column 2), and the total number of schools for which we have value-added for each outcome *after removing outliers* (column 3). The outcome with the most remaining schools is teen arrests, with 127 schools and the least is teen motherhood with 103 (note that this relatively low number is mechanical as this outcome is missing for the 17 all-boys schools).

Table C1. Schools by outcome

	Non-private schools		
	(1)	(2)	(3)
High-Stakes Index	134	126	126
Low-Stakes Index	134	124	124
No Dropout by 14	134	127	124
Not arrested by 18	134	130	127
No live birth by 19	115	103	103
Formally employed 27+	117	111	108

Notes: We show the total number of schools available for each outcome. The first column shows the total number of non-private schools in the cohorts where we have information for each outcome. In the second column we show the remaining amount of schools once we take out those with low first stages. We estimate value-added for all these. Finally, we take out outliers from our estimates and show the remaining number of schools in the last column.

Appendix D: Testing the Validity of the Difference in Difference Variation Across Cutoffs

Robustness to Interactions

Our identification relies that there are no considerable interaction effects between test scores and choices. In order to test for this, we estimate our base model defined in (5) and (6) with four different interactions between choices and SEA scores. The *first interaction* considers the selectivity of choices. For this, we include an interaction between the polynomial of the student's own SEA score and the selectivity of each of the first 4 choices, this being approximated by the average peer SEA score of those assigned to that choice.

$$I_{i,j} = \Sigma(I_{i,\tau} \cdot \pi_{\tau j}) + f_j(SEA_i) + \Sigma(g_{j,k}(SEA_i) \cdot \overline{SEA}_{choice_k}) + \lambda_{j,c} + \mathbf{X}'_{it} \delta_j + S_{j,t} + v_{ijct}, \quad \text{for each } j \in J \quad (18)$$

$$Y_{ijct} = \Sigma(\hat{I}_{i,j} \cdot \theta_j^{TOTIV}) + f(SEA_i) + \Sigma(g_k(SEA_i) \cdot \overline{SEA}_{choice_k}) + \lambda_c + \mathbf{X}'_{it} \delta + S_t + \varepsilon_{ijct} \quad (19)$$

Where $g_k(SEA_i)$ is a fifth order polynomial and $\overline{SEA}_{choice_k}$ is the mean total SEA scores for incoming assigned students for each choice k .

The *second interaction* is between SEA scores and student's choice sets fixed effects. Instead of using one fixed effect per choice set, we use two, one for those below the median SEA score in that choice set and one for those above.

$$I_{i,j} = \Sigma(I_{i,\tau} \cdot \pi_{\tau j}) + f_j(SEA_i) + \lambda_{j,ci} + \mathbf{X}'_{it} \delta_j + S_{j,t} + v_{ijct}, \quad \text{for each } j \in J \quad (20)$$

$$Y_{ijct} = \Sigma(\hat{I}_{i,j} \cdot \theta_j^{TOTIV}) + f(SEA_i) + \lambda_{ci} + \mathbf{X}'_{it} \delta + S_t + \varepsilon_{ijct} \quad (21)$$

Now, instead of λ_c we will have λ_{ci} which will take one value if student i with choicesset c is above the median SEA score within that choice set and another if it is below.

The *third interaction* is between the student's SEA score and its n -th choice of school:

$$I_{i,j} = \Sigma(I_{i,\tau} \cdot \pi_{\tau j}) + f_j(SEA_i) + \Sigma(SEA_i \cdot I_{j,i,k}) + \lambda_{j,c} + \mathbf{X}'_{it} \delta_j + S_{j,t} + v_{ijct}, \quad \text{for each } j \in J \quad (22)$$

$$Y_{ijct} = \Sigma(\hat{I}_{i,j} \cdot \theta_j^{TOTIV}) + f(SEA_i) + \Sigma(SEA_i \cdot I_{i,k}) + \lambda_c + \mathbf{X}'_{it} \delta + S_t + \varepsilon_{ijct} \quad (23)$$

Where $I_{i,k}$ is an indicator equal to 1 if the n -th choice of student i is school k . This is done for the first 4 choices.

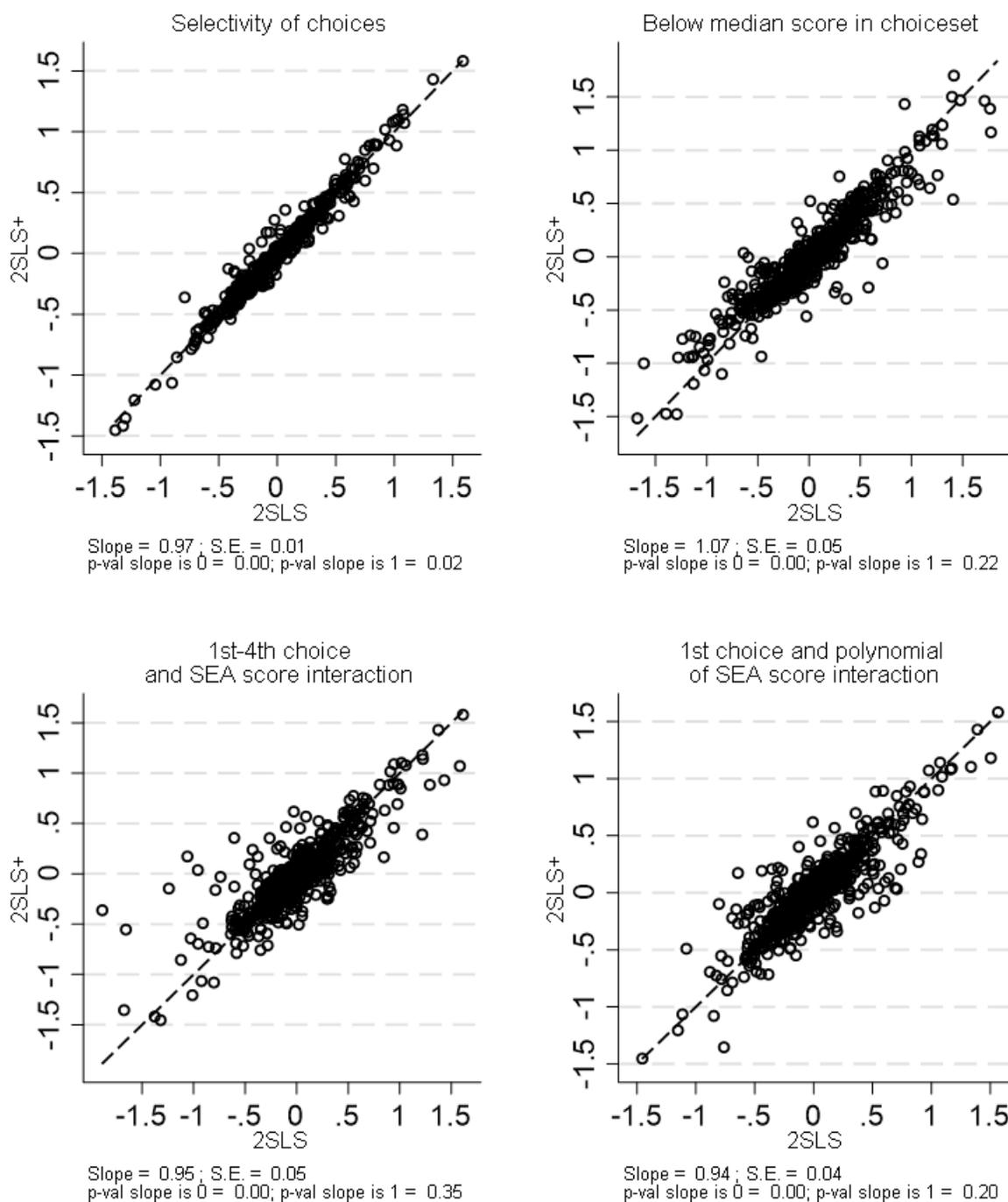
Finally, the *fourth interaction* is between the student's SEA score polynomial and its 1st choice of school:

$$I_{i,j} = \Sigma(I_{i,\tau} \cdot \pi_{\tau j}) + f_j(SEA_i) + \Sigma(g_{j,k}(SEA_i) \cdot I_{j,i,k}) + \lambda_{j,c} + \mathbf{X}'_{it} \delta_j + S_{j,t} + v_{ijct}, \quad \text{for each } j \in J \quad (24)$$

$$Y_{ijct} = \Sigma(\hat{I}_{i,j} \cdot \theta_j^{TOTIV}) + f(SEA_i) + \Sigma(g_k(SEA_i) \cdot I_{i,k}) + \lambda_c + \mathbf{X}'_{it} \delta + S_t + \varepsilon_{ijct} \quad (25)$$

Where $g_k(SEA_i)$ is a fifth order polynomial and $I_{i,k}$ is an indicator equal to 1 if the first choice of student i is school k . Results for all four models are shown in [Figure D1](#)

Figure D1: Robustness of School Impacts to Interactions Between Choices and Incoming Scores



Notes: We show the relationship between the estimated school impacts with the interacted model (y-axis) and the model without interactions (x-axis). Each panel shows results for one of the interacted models. Estimated slope and p-values resulting from testing for whether the slope differs from both 0 and 1 are shown below the graph. The short dashed line shows the 45° relationship.

Regression Discontinuity Variation vs. All Variation

Existing papers that have explored parental preferences for school causal impacts have either relied on school average outcomes (which may not reflect their impacts *per se*) or estimated school impacts that may be biased due to selection.⁵⁸ If one's measures of school effectiveness do not accurately reflect schools' causal impacts, it may distort one's conclusions regarding parental preferences for school effectiveness. For this reason, validating the estimated school impacts as reflecting *causal* impacts is important. A key strength of our context and data is that we are able to validate our estimated school impacts using exogenous variation only. We test the validity of our value-added estimates by exploring if they are consistent with what one would obtain using plausibly exogenous variation only.

Under the algorithm used to create the tentative school assignments (discussed in [Appendix A](#)), each school has a minimum score above which applicants are tentatively admitted and below which they are not. As such, the marginal effect of being tentatively assigned to each school (relative to the next lowest ranked school) can be estimated with a regression discontinuity design. That is, among students who are applicants to a given school τ , the causal effect of being tentatively assigned to school τ is simply the effect of scoring above the admission cutoff for school τ (conditional on smooth functions of ones incoming SEA score). In our setup, students are considered to be applicants to a school if that school is in their ranked list and they do not score above the cutoff for a more preferred school. Note, therefore, that students can be applicants to more than one school.⁵⁹

To obtain the reduced-form Regression Discontinuity (RD) effect of being tentatively assigned to any school τ , we estimate RD models for each outcome among all applicants to school τ .⁶⁰ Under the RD identifying assumptions, the reduced-form effect of being tentatively assigned to school τ on outcome Y , is captured by estimating the equation below.

$$Y_{ij} = Above_{i\tau} \cdot \gamma_{\tau} + f(SEA_i) + \mathbf{X}'_i \delta + \varepsilon_{ij} \quad (26)$$

Where Y_{ij} is the outcome of student i who attended school j , and $Above_{i\tau}$ is an indicator for scoring above the algorithm-based assignment cutoff for school τ . Among those who comply with the cutoff, $j=\tau$. The parameter γ_{τ} captures the difference in outcomes (all else equal) between those exogenously assigned to a preferred school τ (due to scoring above the cutoff) versus scoring below the cut off and attending the student's counterfactual school q (that is, the school that the students would have attended had they not scored above the cutoff for school τ). As such, in the neighborhood of the cutoff,

$$E[\hat{\gamma}_{\tau} | X_i, SEA_i] = E(Y_{ij} | Above = 1) - E(Y_{iq} | Above = 0) \quad (27)$$

To simplify equation (27), we consider this expression for compliers and for non-compliers.

⁵⁸In related work [Abdulkadiroğlu et al. \(2020\)](#) examine parents responsiveness to school impacts that rely on selection on observables assumptions (similar to our estimates). However, they are unable to validate these school impact estimates using exogenous variation in school attendance.

⁵⁹For example, a student that was assigned to her first choice will only appear once as a (successful) applicant to her top choice school. However, a student who is assigned to her second choice school will appear twice: as an (unsuccessful) applicant to her top choice school and as a (successful) applicant to her second choice school.

⁶⁰That is we estimate separate reduced-form models where, in each one, we consider all persons who applied to a particular school τ in each year.

Under the assumption of unconfoundedness (Rubin 1990), it follows that $E[Y_{ij} - Y_{iq}|X_i, SEA_i] = \theta_j^{TOT} - \theta_q^{TOT}$. That is, if there is no selection on observables, the average difference in outcomes between observationally equivalent individuals who attended school j and school q reflects the difference in causal impacts between school j and school q . Among the compliers, school j is school τ if they score above the cutoff.⁶¹ As such, for compliers, $E[\hat{\gamma}_\tau|X_i, SEA_i] = \theta_\tau^{TOT} - E[\theta_q^{TOT}]$, where $E[\theta_q^{TOT}]$ is the average impact of the counterfactual schools for the applicants to school τ . Among non-compliers, the cutoff does not change the school attended so that $E[\hat{\gamma}_\tau|X_i, SEA_i] = 0$. It follows that for the average applicant to school τ , equation (27) can be written as equation (28) below.

$$E[\hat{\gamma}_\tau|X_i, SEA_i] = \bar{p}_\tau \times (\theta_\tau^{TOT} - E[\theta_q^{TOT}]) \quad (28)$$

In words, in expectation, the estimated effect of scoring above the cutoff for school τ is the difference between the impact of attending preferred school τ and that of attending the average counterfactual school q , all times the compliance rate (\bar{p}_τ). This is simply the weighted cutoff effect for the compliers and the non-compliers.

Consider now, estimating this same model, but replacing each student's actual outcome with the predicted TOT impact of the school they attended, $\hat{\theta}_j^{TOTIV}$, as below.

$$\hat{\theta}_j^{TOTIV} = Above_{i\tau} \cdot \zeta_\tau + f(SEA_i) + \mathbf{X}_i' \delta + \varepsilon_{ij} \quad (29)$$

The parameter ζ_τ is the difference in predicted TOT school impacts (all else equal) between those scoring above the cutoff for preferred school τ versus not. In the neighborhood of the cutoff, the RD effect on the predicted TOT impacts of an individual's attended school is $E[\zeta_\tau|X_i, SEA_i] = E(\hat{\theta}_j^{TOTIV} | Above = 1) - E(\hat{\theta}_q^{TOTIV} | Above = 0)$. Using the same logic as above for compliers and non-compliers, it follows that

$$E[\hat{\zeta}_\tau|X_i, SEA_i] = \bar{p}_\tau \times (E[\hat{\theta}_\tau^{TOTIV}] - E[\hat{\theta}_q^{TOTIV}]) \quad (30)$$

In words, in expectation, the estimated difference in predicted school TOT impacts of scoring above the cutoff for school τ is the difference between the estimated TOT impact of attending preferred school τ and that of the average counterfactual school q , all times the compliance rate (\bar{p}_τ).

Inspection of (28) and (30) reveals that if our treatment on the treated estimated impacts for attended school j and the average counterfactual school q are unbiased, then by the law of iterated expectations, $E[\hat{\zeta}_\tau|X_i, SEA_i] = E[\hat{\gamma}_\tau|X_i, SEA_i]$. In words, if the estimated TOT school impacts ($\hat{\theta}_\tau^{TOTIV}$) are consistent estimates of the causal effect of attending school τ , then for each applicant school, the RD estimates using the actual outcomes as left hand side variables and the RD estimates using the TOT school impacts of the attended school j as left hand side variables, should be equal in expectation.

This motivates a validation test of our TOT school estimates. In related work, Hastings et al. (2015) implement a very similar test to validate the reliability of using predicted versus actual earnings when disseminating information on the expected returns to attend alternative colleges and majors. To implement this test, first, we estimate $\hat{\gamma}_\tau$ and $\hat{\zeta}_\tau$ for each preferred school τ (using the optimal bandwidth from Imbens and Kalyanaraman (2012)). Following Hastings et al. (2015), to

⁶¹In most instances school q will be the next ranked school in the choice list, but could be any fallback school (such as a private school or Government Assisted school that can admit students irrespective of their SEA scores).

account for estimation errors in the RD effects on school impacts, we implement Empirical Bayes estimates of each cutoff effect.⁶² We then regress the former estimated coefficients on the latter (to account for noisiness in the RD estimated effects, we weight each estimate by the inverse of its squared standard error when doing this regression). Finally, we test for whether the estimated slope is statistically indistinguishable from 1. The results from this approach are reported in Figure 2 in the main text.

⁶²Specifically, for any particular outcome, the *predicted* RD effect for school j is the weighted difference in estimated school impacts between those just above and below the cutoff for school j . We can express the estimated parameter as $\hat{\zeta}_{\hat{\theta}_j} = \sum_{k \in A} a_k(\hat{\theta}_k) - \sum_{k \in B} b_k(\hat{\theta}_k) = \sum_{k \in A} a_k(\theta_k + \varepsilon_k) - \sum_{k \in B} b_k(\theta_k + \varepsilon_k)$, where A is the set of schools the students attend above the cutoff, B the set of schools the students attend below the cutoff, a_k and b_k the proportions in which they do so and ε_k the estimation error for the school impact θ_k . If we assume that the school impacts are not independent within each cutoff, but that the estimation errors are, we can approximate the variance of this estimated parameter as $Var(\hat{\zeta}_{\hat{\theta}_j}) = \sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta} + \sum w_k SE_k^2$, where σ_{θ}^2 is given by the magnitude of the school impacts, Cov_{θ} is approximated by the covariance between each pair of schools the students applied to and SE_k^2 is given by the square of the standard error of the school impact. If we also include the squared standard error of the RD estimate, $SE_{\hat{\zeta}}$, the reliability ratio of our RD estimate is given by

$$\lambda_j = \frac{\sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta}}{(\sum u_k \sigma_{\theta}^2 + \sum v_{lm} Cov_{\theta}) + (\sum w_k SE_k^2 + SE_{\hat{\zeta}}^2)} \quad (31)$$

Our Empirical Bayes estimate of the *predicted* effect of cutoff j is therefore $[\lambda_j \times \hat{\zeta}_{\hat{\theta}_j}]$

Appendix E: Influence of Admission Probabilities on School Choices

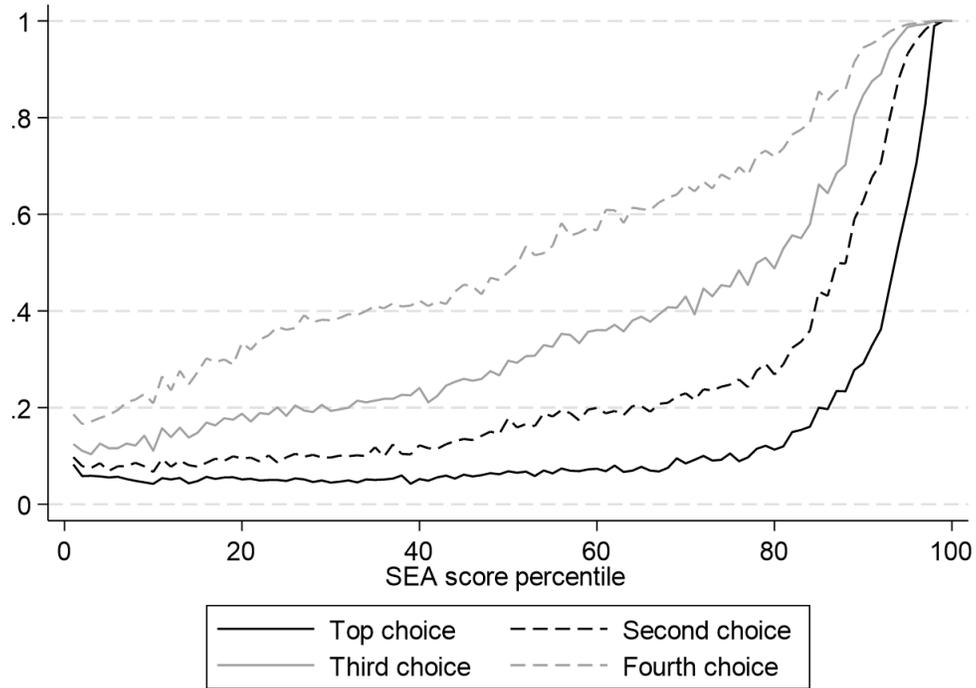
As discussed in Section V.3, a key conditioning variable in our analysis is the admission probability. Figure E1 plots admission probabilities for all the range of SEA scores differentiated by the priorities of the submitted school choices. Consistent with rational behavior, students rank more selective schools (where the probability of accessing them is almost null for students below the 80th percentile of the incoming test score distribution) higher within their choices.

Figure E2 shows the coefficients on the rejection probabilities from a choice model that includes these probabilities (only in first pseudo-observation) alone but without interacting these probabilities with the school's causal impacts or average outcomes.⁶³

More desirable schools (for both observed and unobserved reasons) are those with higher rejection probabilities, by construction. As such, positive coefficients on rejection probabilities would indicate that there are unaccounted-for school attributes that are positively correlated with the rejection probabilities. However, if the school attributes included accurately reflect those dimensions of school quality that parents care for, the coefficients on the rejection probabilities should be negative. In the impacts only model, the point estimates for all groups are negative (Figure D2, left) – that is, conditional on peer quality, proximity and impacts on key outcomes, parents are less likely to choose schools to which their child is less likely to be admitted. In the full model (Figure D2, right), as expected, the coefficients on the rejection probabilities become more negative – indicating that schools that are more desirable tend to have better average outcomes. The fact that the coefficients on the rejection probabilities are negative for all groups provides empirical validation of the theoretical predictions from Chade and Smith (2006), and suggests that our included variables appropriately capture relevant determinants of parents' school choices.

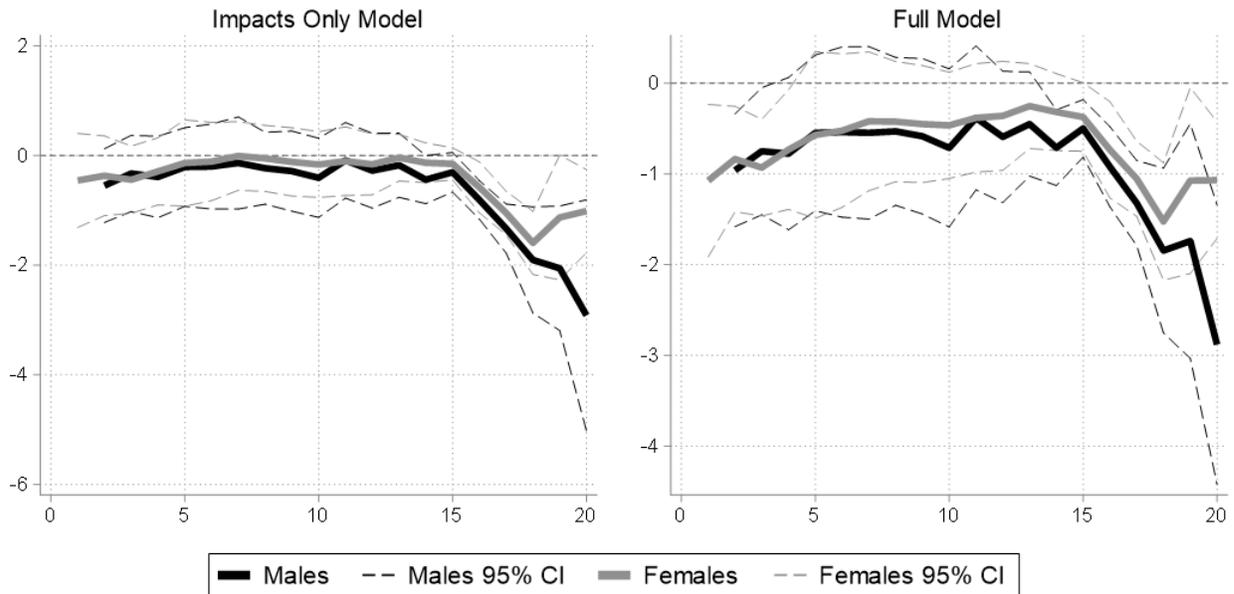
⁶³We estimate this model to provide a more easily understandable relation between school choices and rejection probabilities (as opposed to the main model with interactions where the estimated coefficients on the rejection probabilities are not directly interpretable).

Figure E1. Likelihood of getting into each choice by SEA score percentile



Notes: The X-axis represents the SEA score percentile. The Y-axis represents the likelihood of being assigned to a given choice. This probability is higher for the top SEA percentiles. The top choice, which are usually more selective schools, have higher cutoffs and a lower probability of assignment for each SEA percentile in the distribution. As we move to the next preferred choices, the probability of being assigned to that given choice increases for all SEA percentiles.

Figure E2: Rejection Probability

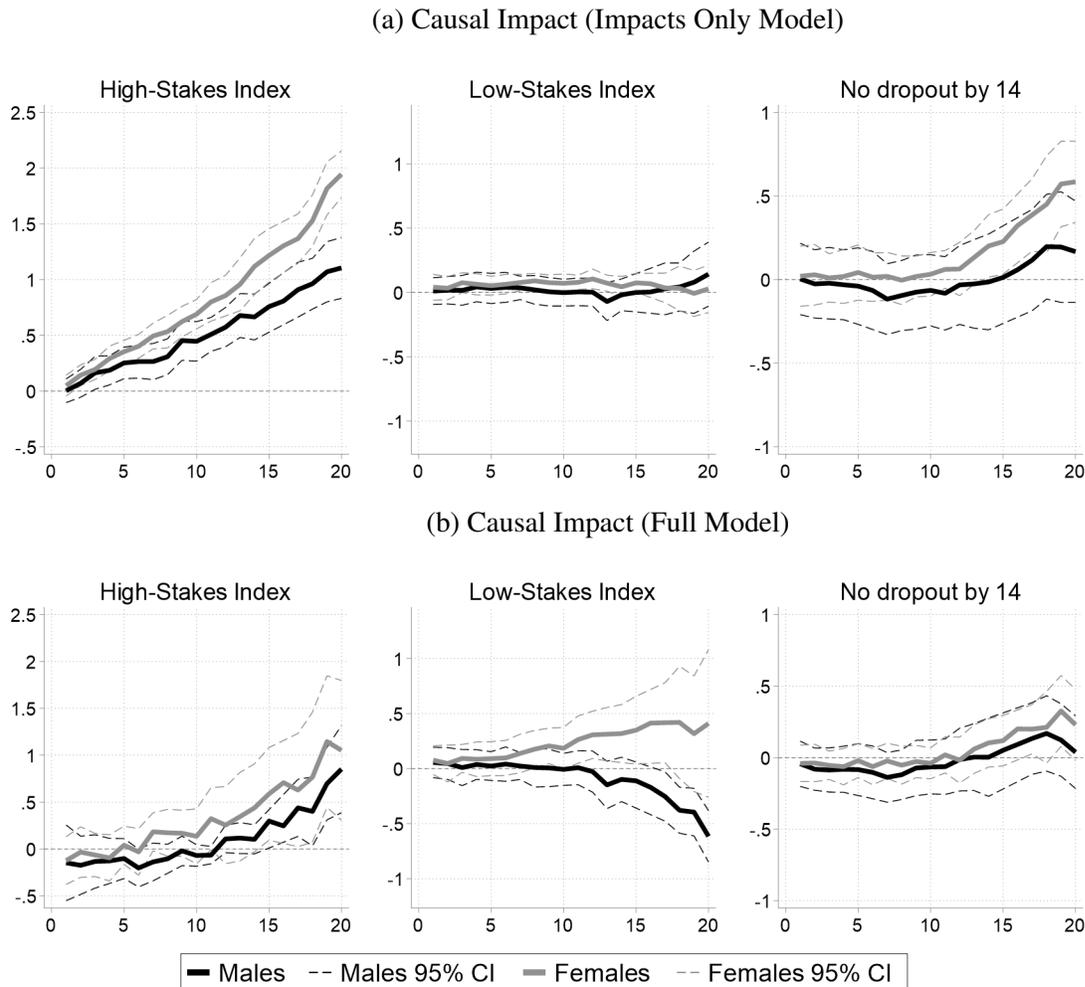


Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: The left panel displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and the right panel displays estimates from the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, and the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools. The dashed lines represent the associated 95% confidence intervals.

Appendix F: Choice Model using Alternative Specifications

Rank-Ordered Logit: We first present estimates from this standard model which neither includes the first pseudo-observation nor the rejection probabilities. See Figures F1-F3.

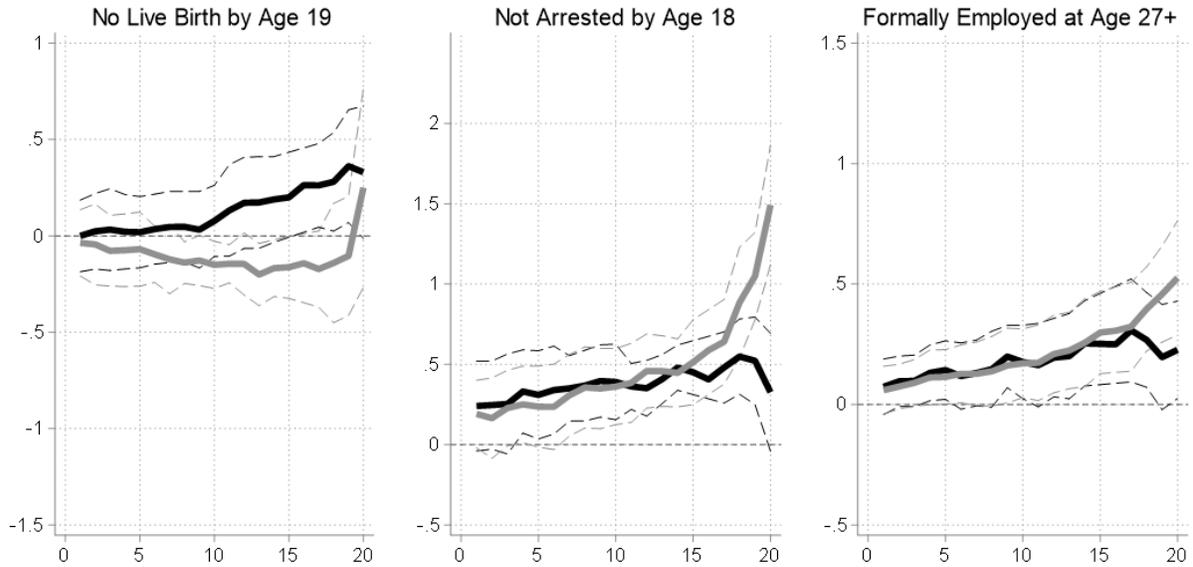
Figure F1. Academic Outcomes using Rank-Ordered Logit Model



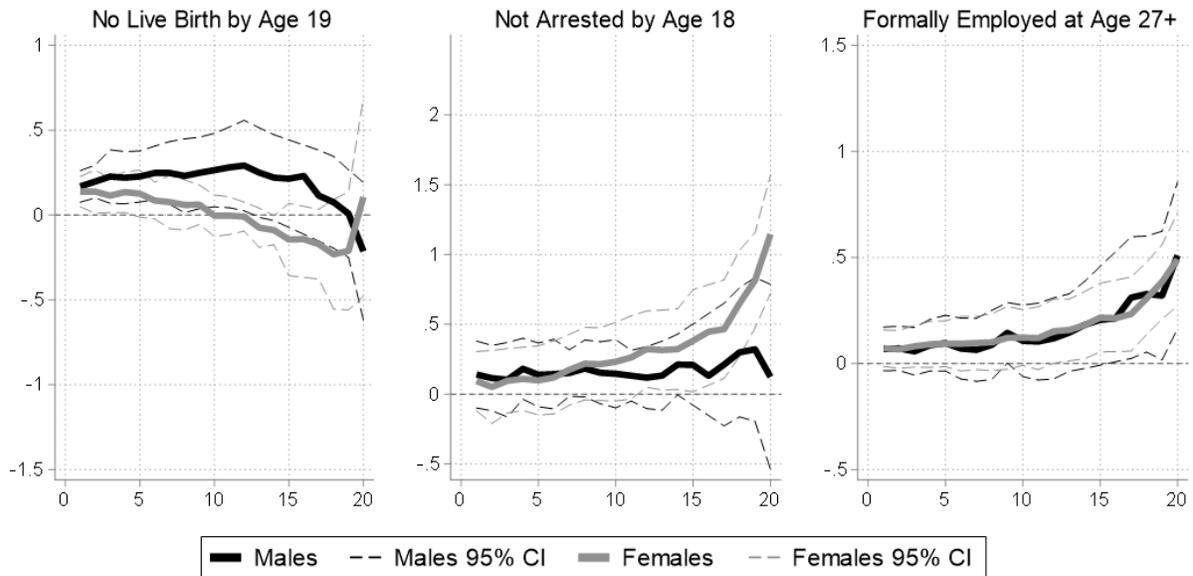
Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Rank Ordered Logit Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Rank Ordered Logit Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, only comparing ranked schools. The dashed lines represent the associated 95% confidence intervals.

Figure F2. Non-Academic Outcomes using Rank-Ordered Logit Model

(a) Causal Impact (Impacts Only Model)

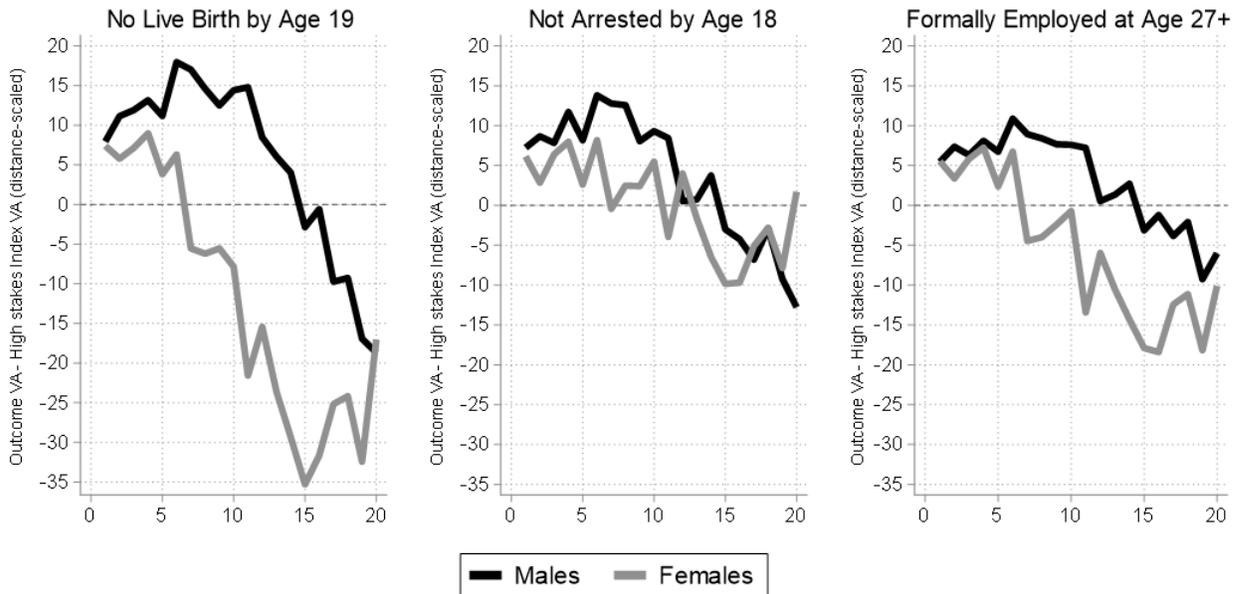


(b) Causal Impact (Full Model)



Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile)×(gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Rank Ordered Logit Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Rank Ordered Logit Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, only comparing ranked schools. The dashed lines represent the associated 95% confidence intervals.

Figure F3. Comparison of Choice Model's Estimated Coefficients (Rank-Ordered Logit Model)

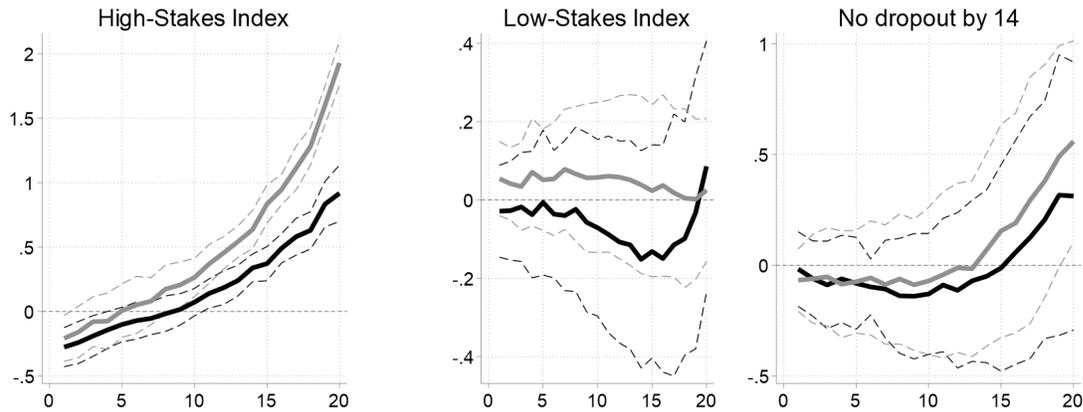


Notes: This figure presents the difference between the choice model's estimates on the school impacts of three non-academic outcomes and the choice model's estimate on the school impacts of the high-stakes index, scaled by the log distance estimate. The X-axis represents the individual score ventile. The connected lines represent the difference between the choice model estimate on the non-academic impacts and the choice model estimate on the high-stakes impacts divided by the log distance cell estimate (and scaled by 6 for ease of interpretation). This difference is computed separately for each (SEA score ventile) × (gender) cell. The estimates result the Ranked Ordered Logit Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality and log distance, control variables (whether the secondary school is on the same island, whether it is all-girls, whether it is all boys) only for the ranked schools.

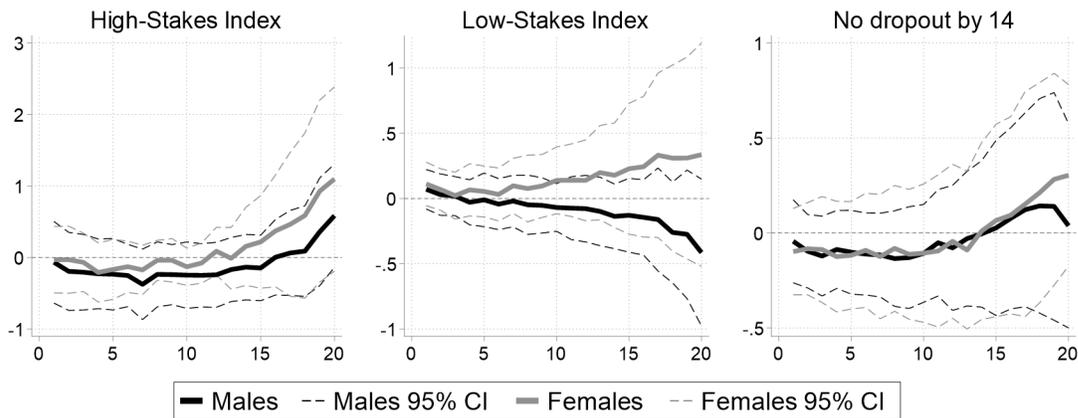
No Rejection Probabilities: We also present results from a model that includes the first pseudo-observation but does not account for rejection probabilities in any form. See Figures F4-F6.

Figure F4. Academic Outcomes using The Model without Rejection Probabilities

(a) Causal Impact (Impacts Only Model)



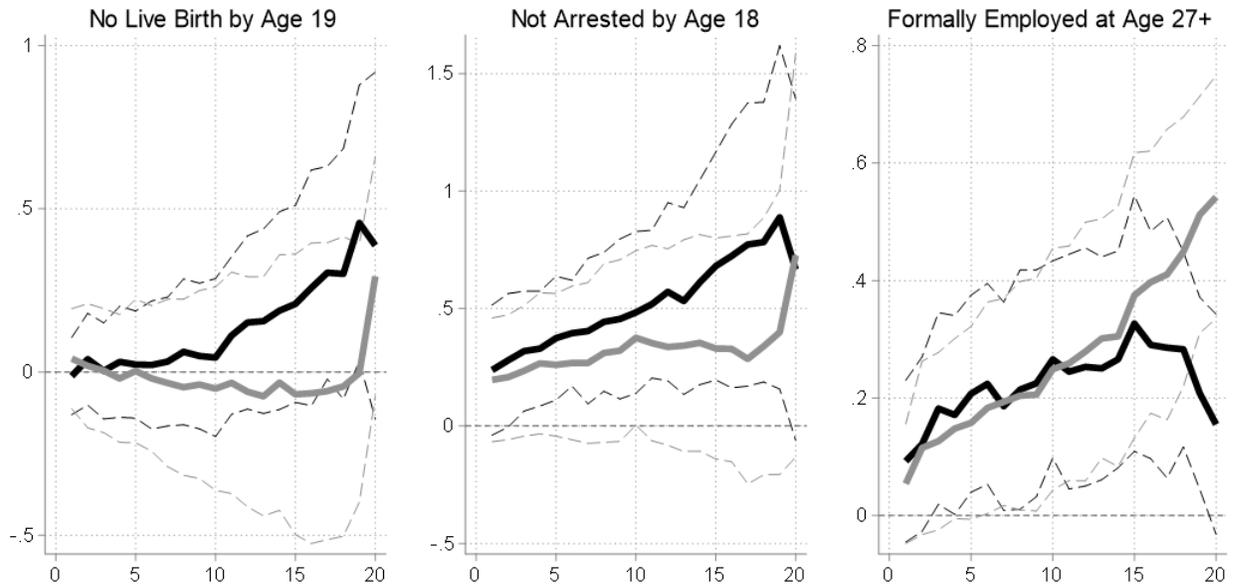
(b) Causal Impact (Full Model)



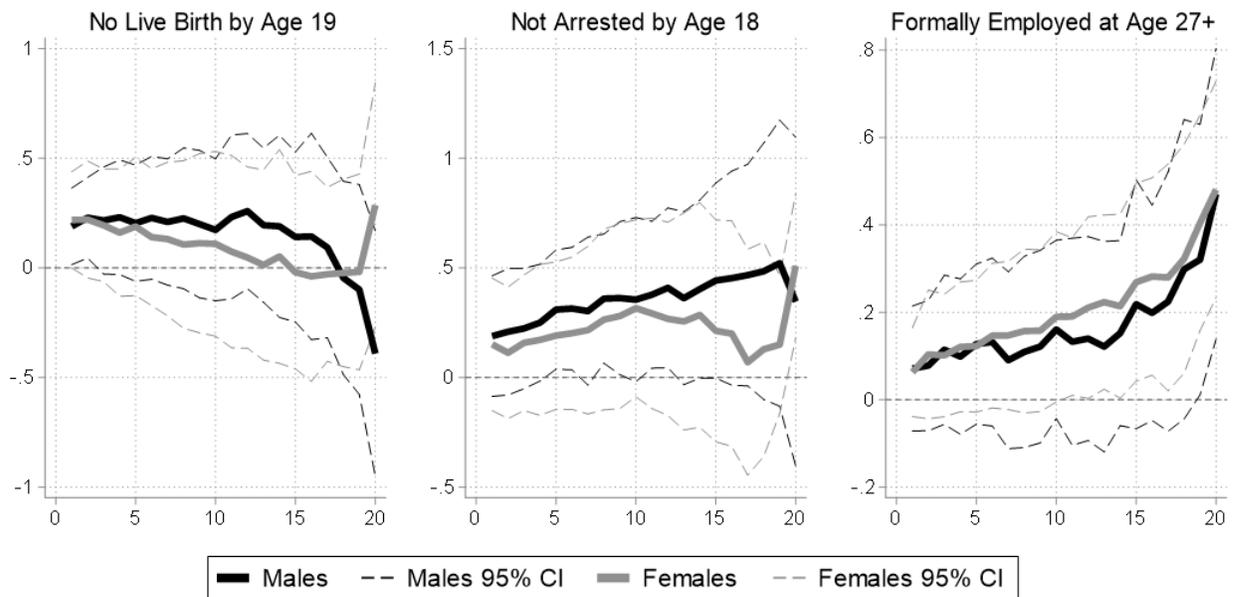
Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model using the model without rejection probabilities, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model using the model without rejection probabilities, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, and whether it is all-boys. The dashed lines represent the associated 95% confidence intervals.

Figure F5. Non-Academic Outcomes using The Model without Rejection Probabilities

(a) Causal Impact (Impacts Only Model)

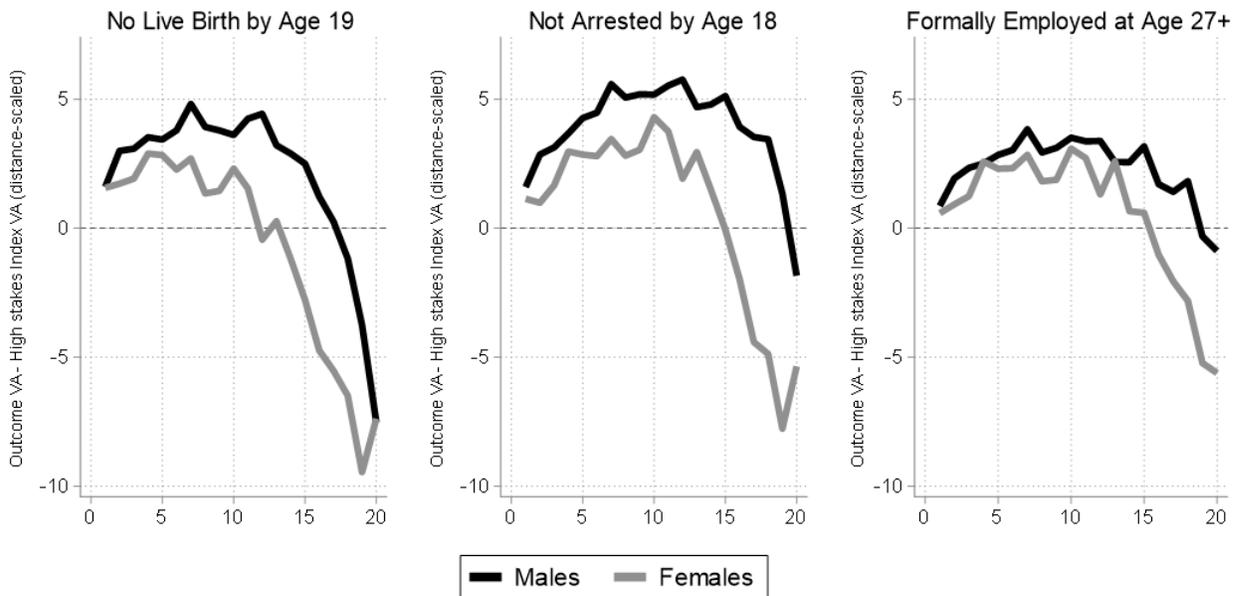


(b) Causal Impact (Full Model)



Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile)×(gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model using the model without rejection probabilities, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model using the model without rejection probabilities, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, and whether it is all-boys. The dashed lines represent the associated 95% confidence intervals.

Figure F6. Comparison of Choice Model's Estimated Coefficients (Model without Rejection Probabilities)

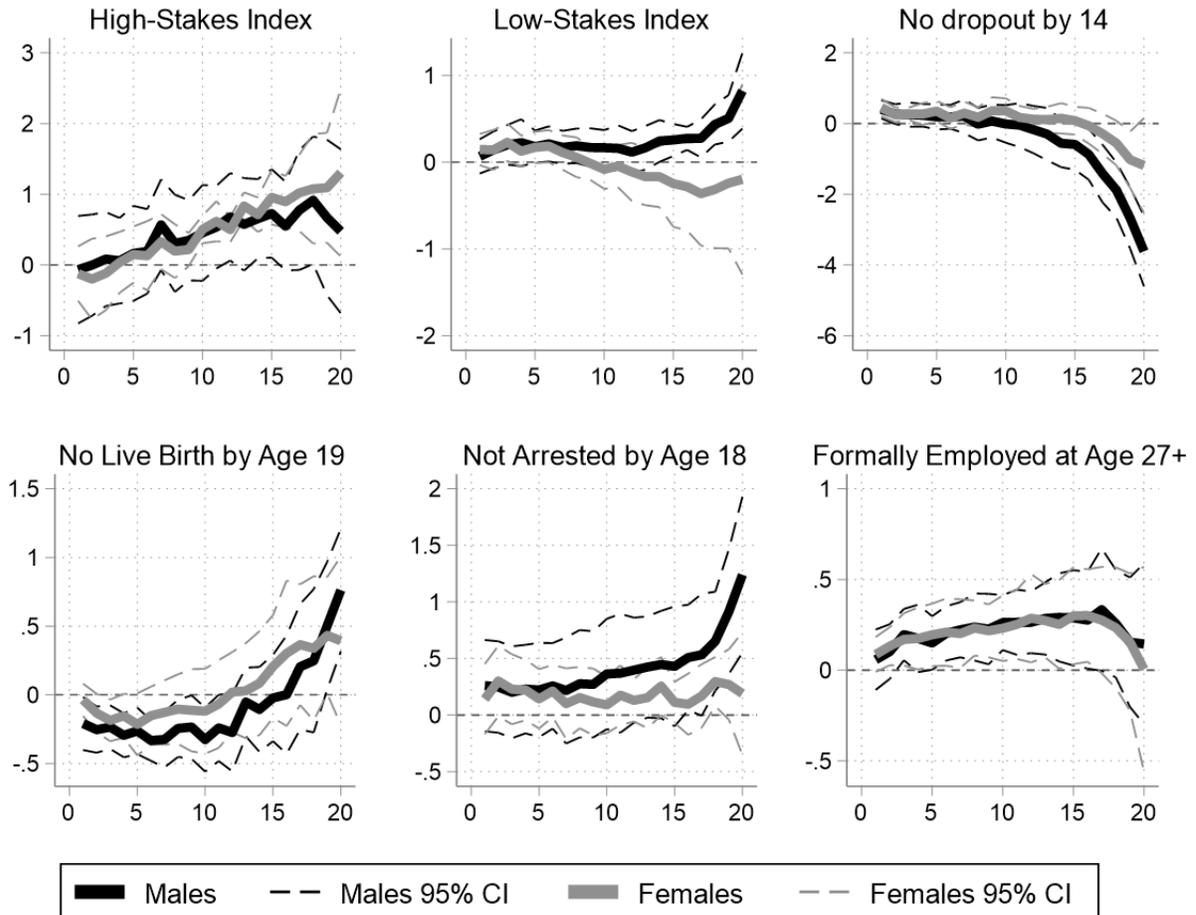


Notes: This figure presents the difference between the choice model's estimates on the school impacts of three non-academic outcomes and the choice model's estimate on the school impacts of the high-stakes index, scaled by the log distance estimate. The X-axis represents the individual score ventile. The connected lines represent the difference between the choice model estimate on the non-academic impacts and the choice model estimate on the high-stakes impacts divided by the log distance cell estimate (and scaled by 6 for ease of interpretation). This difference is computed separately for each (SEA score ventile) × (gender) cell. The estimates result from the Full Model (but without the inclusion of rejection probabilities), which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality and log distance, control variables (whether the secondary school is on the same island, whether it is all-girls, whether it is all boys).

Appendix G: School-Level Average Outcomes

This appendix plots the estimated coefficients on the school average outcomes from the Full Model described in the text, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality and log distance, control variables (whether the secondary school is on the same island, whether it is all-girls, whether it is all boys).

Figure G1. Average Outcomes (Full Model)

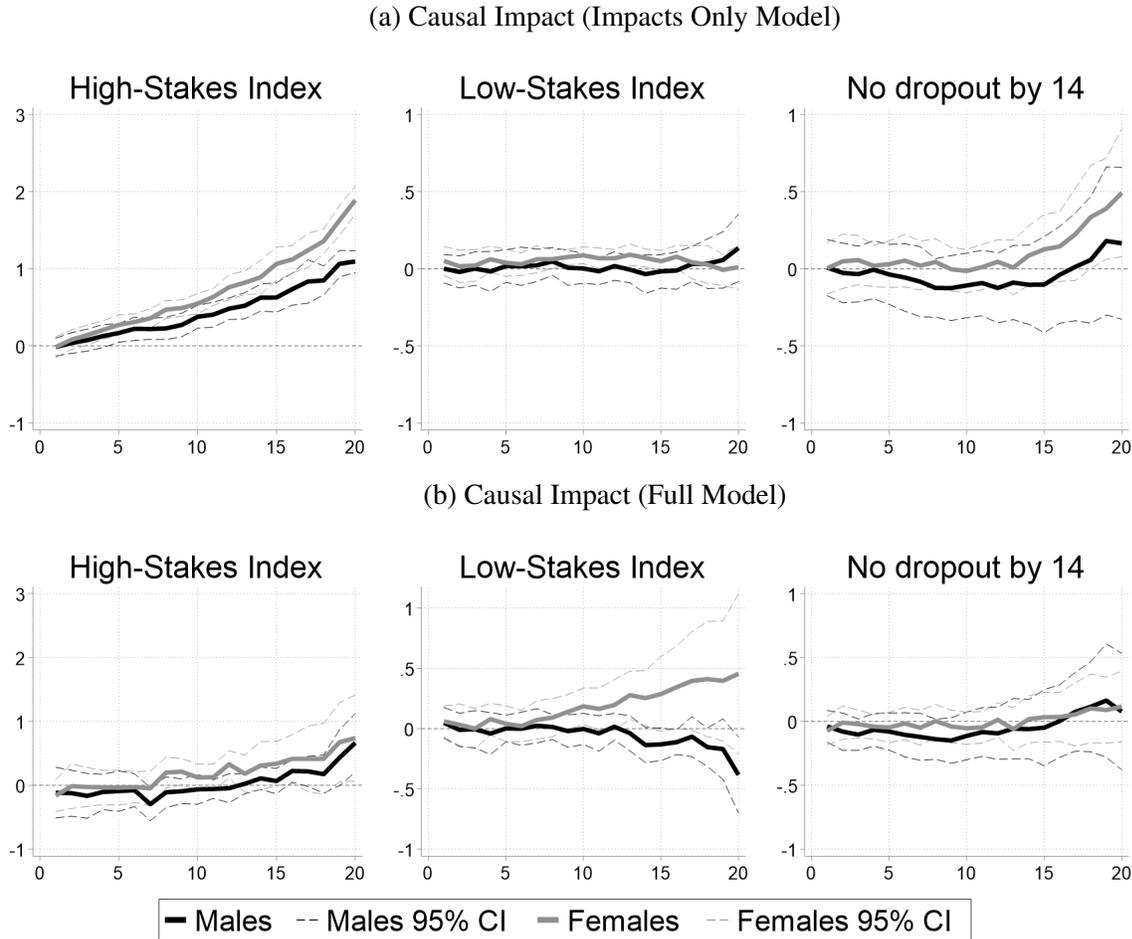


Notes: The X-axis represents the SEA score ventile. The connected lines represent the estimated coefficients of the standalone average outcomes, computed separately for each (SEA score ventile)×(gender) cell, for the Full Model, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Appendix H: Choice Model with Out of Sample Impacts and Averages

This appendix reports estimates from the same choice models described in the text but using leave-year-out school impacts and average outcomes. That is, for each SEA cohort, we estimate schools' causal impacts and compute average outcomes disregarding the information of the cohort for which these calculations are being implemented.

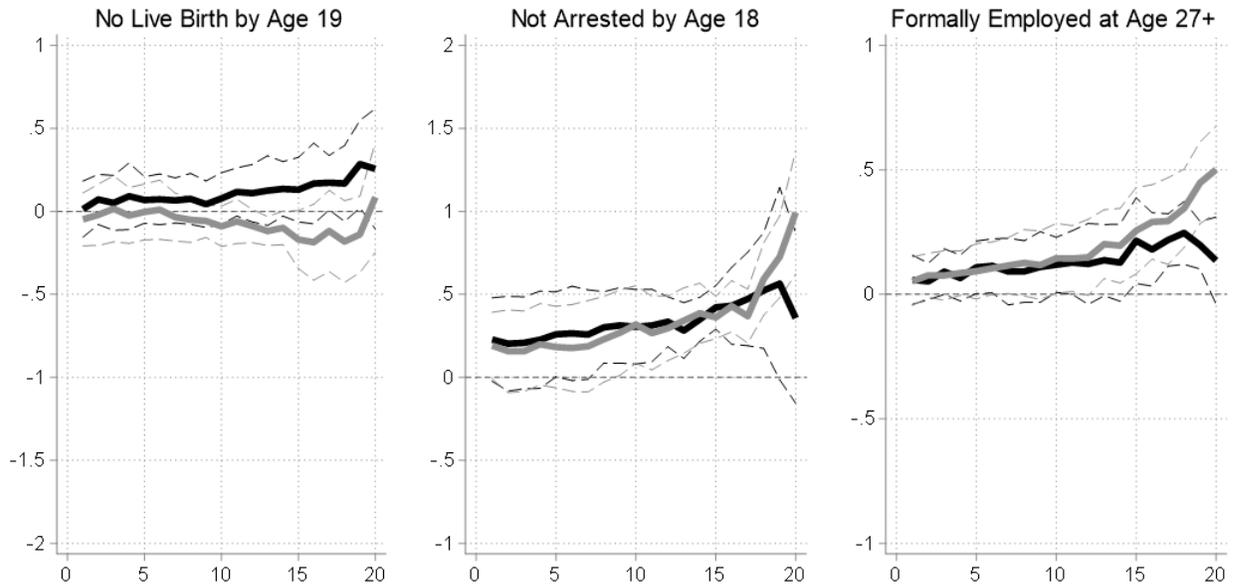
Figure H1. Academic Outcomes using Out of Sample Estimated Impacts and Average Outcomes



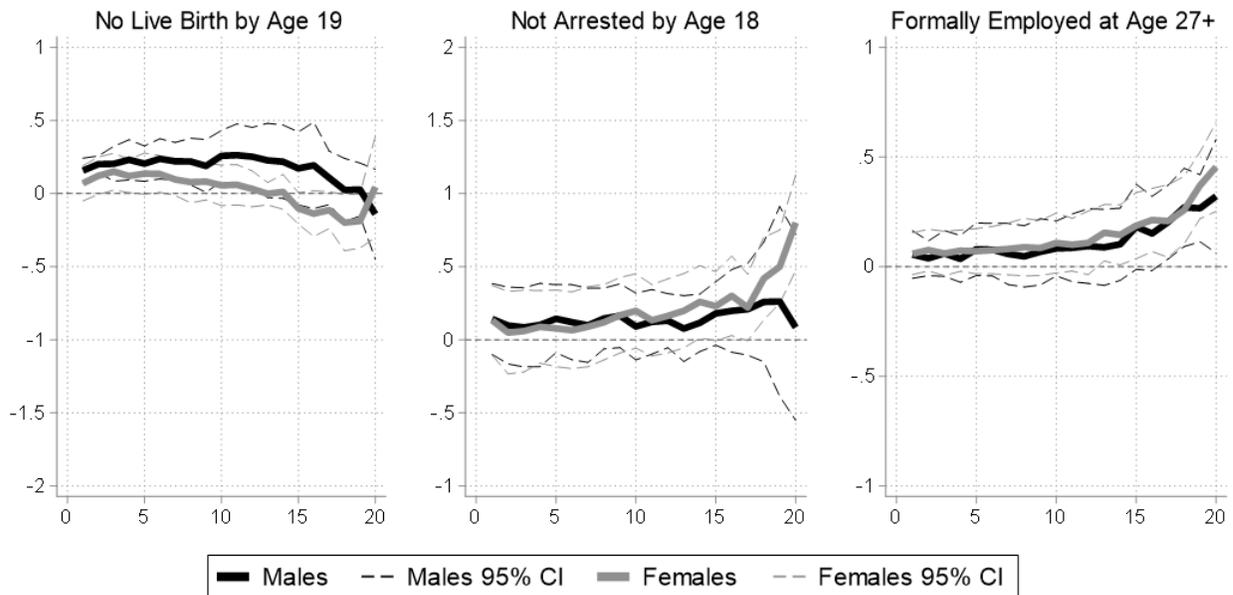
Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile)×(gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model using out of sample estimates of schools' causal impacts for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model using out of sample estimates of schools' causal impacts for all outcomes, out of sample school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Figure H2. Non-Academic Outcomes using Out of Sample Estimated Impacts and Average Outcomes

(a) Causal Impact (Impacts Only Model)

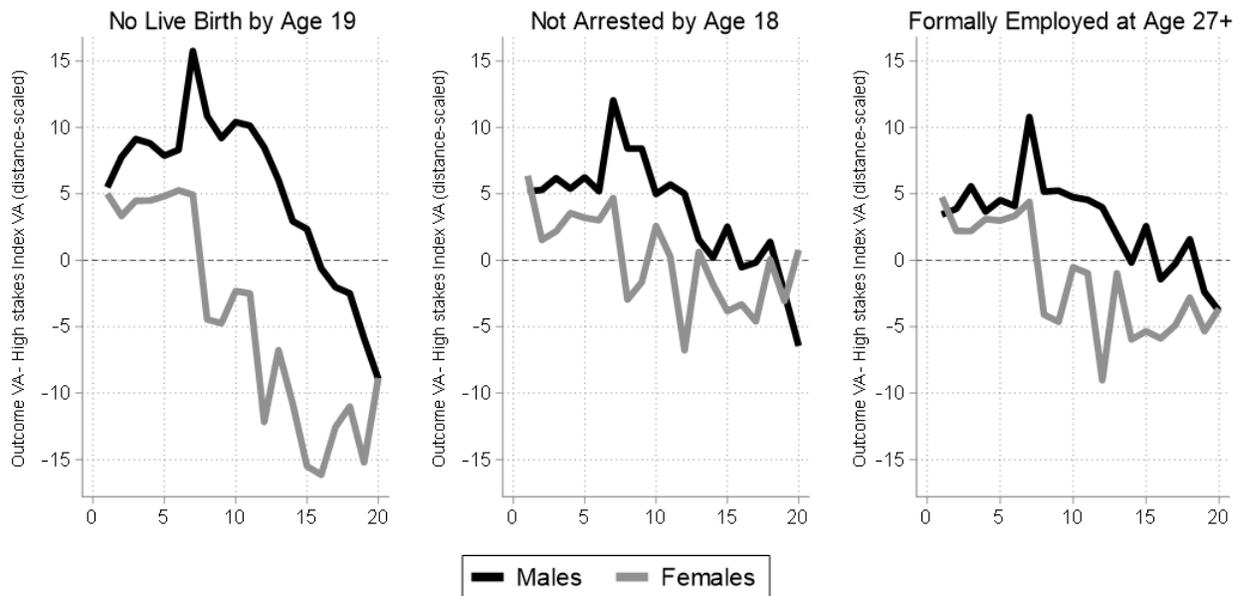


(b) Causal Impact (Full Model)



Notes: The X-axis represent the individual SEA score ventile. The connected lines represent the estimated coefficients, computed separately for each (SEA score ventile) \times (gender) cell, for two main models: Panel (a) displays estimates from the Impacts Only Model using out of sample estimates of schools' causal impacts for all outcomes, peer quality, and log distance and Panel (b) displays estimates from the Full Model using out of sample estimates of schools' causal impacts for all outcomes, out of sample school-level averages for all outcomes, peer quality, and log distance. All specifications include control variables for whether the secondary school is on the same island, whether it is all-girls, whether it is all-boys, the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes. The dashed lines represent the associated 95% confidence intervals.

Figure H3. Comparison of Choice Model's Estimated Coefficients (Out of Sample Estimated Impacts and Average Outcomes)



Notes: This figure presents the difference between the choice model's estimates on the school impacts of three non-academic outcomes and the choice model's estimate on the school impacts of the high-stakes index, scaled by the log distance estimate. The X-axis represents the individual score ventile. The connected lines represent the difference between the choice model estimate on the non-academic impacts and the choice model estimate on the high-stakes impacts divided by the log distance cell estimate (and scaled by 6 for ease of interpretation). This difference is computed separately for each (SEA score ventile)×(gender) cell. The estimates result from the Full Model using out of sample estimates, which includes schools' causal impact estimates for all outcomes, the school-level averages for all outcomes, peer quality and log distance, control variables (whether the secondary school is on the same island, whether it is all-girls, whether it is all boys), the estimated likelihoods of school rejection only when comparing the top choice with all unranked schools, and interactions of these likelihoods with all school attributes.