

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

DIETHER W. BEUERMANN

Inter-American Development Bank

C. KIRABO JACKSON

Northwestern University

LAIA NAVARRO-SOLA

Stockholm University

and

FRANCISCO PARDO

University of Texas at Austin

First version received September 2020; Editorial decision January 2022; Accepted April 2022 (Eds.)

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 labour 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 labour 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.

Key words: School Value-Added, School Choice, School Preferences, Trinidad and Tobago

JEL Codes: I20, J0

1. 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 labour 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 article are twofold. First, we aim to better understand the multidimensional nature of individual schools' output. Researchers, practitioners, and policy-makers often rely on schools' performance on standardized tests as a measure of quality. However, because educational output may be multidimensional (Hanushek, 1971; Heckman, Stixrud and Urzua, 2006; Kautz, Heckman, Diris, Weel and Borghans, 2017; Jackson, 2018; Jackson *et al.*, 2020), 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 schools' impacts across several different outcomes.¹ However, to date, only four studies examine the causal impact of individual schools on different outcomes (Abdulkadiroğlu, Angrist and Pathak, 2014; Dobbie and Fryer, 2015; Angrist, Cohodes, Dynarski, Pathak and Walters, 2016; Place and Gleason, 2019). To rely on *causal* impacts, these studies focus on a small number of oversubscribed schools that admit students using enrolment exams or randomized lotteries.² While focusing on oversubscribed schools 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 non-academic outcomes. As such, no studies have used a quasi-experimental design to identify individual schools' 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

1. We 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, Hastings, Kane and Staiger (2014) find that school choice lotteries improve test scores and educational attainment (only for girls). Beuermann and Jackson (2022) find that attending a preferred school in Barbados improves long-run outcomes but not test scores. Also, Booker, Sass, Gill and Zimmer (2011) find that charter school attendance impacts on test scores do not correlate with their impacts on college outcomes. All these studies examine groups of schools rather than individual school impacts—precluding an analysis of the multidimensional nature of educational output by schools.

2. Place and Gleason (2019) and Angrist *et al.* (2016) examine 31 and 26 oversubscribed charter schools, respectively. Abdulkadiroğlu *et al.* (2014) examine six elite selective enrolment schools, and Dobbie and Fryer (2015) examine a single charter school. Dobbie and Fryer (2020) examine impacts of 57 charter schools in Texas but rely on a selection on observables strategy so that (*as they acknowledge*) the estimates may not capture causal impacts.

(98.5% 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, 2022). The few studies that directly examine preferences for school causal impacts conclude that parents may not value school impacts *per se* (Rothstein, 2006; Abdulkadiroğlu, Pathak, Schellenberg and Walters, 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 US. 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 non-academic 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 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 3 years later, high-stakes secondary school completion exams 5 years later, and a national tertiary certification exam taken 7 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 scores on the Secondary Entrance Assessment. 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.

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

Abdulkadiroğlu *et al.*, 2014; Dobbie and Fryer, 2015; Angrist *et al.*, 2016; Place and Gleason, 2019; Deutsch, Gill and Johnson, 2020; Dobbie and Fryer, 2020).

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*, (1) not all school rankings can be observed, and (2) 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 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 artefact of the chosen methodology, our main findings are similar in models that do not account for strategic behaviours (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 labour 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 non-academic 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).

In terms of parental school preferences, we first replicate results of existing studies. Parents assign higher rankings to more proximate schools, higher-performing schools, and those with higher-achieving peers (Burgess, Greaves, Vignoles and Wilson, 2015; Abdulkadiroğlu *et al.*, 2020). However, 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

4. Researchers 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, Kane and Staiger (2009) appeal to patterns in the data to justify the assumption that choices made are not strategic.

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 non-academic outcomes, we find robust evidence that parents prefer schools that reduce arrests, reduce teen births, and increase formal labour 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 non-academic 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 labour 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 labour 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 multi-dimensionality 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 labour market participation. We show that parents may have strong preferences for schools that reduce crime, reduce teen births, and increase labour 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' non-academic output) may be very misleading about their welfare effects.

The remainder of this article is as follows: Section 2 describes the Trinidad and Tobago context and discusses the data. Section 3 presents our empirical strategy for estimating school causal impacts. Section 4 presents the magnitudes of the estimated school impacts and explores the potential multidimensionality of school output. Section 5 discusses our choice models and presents estimates of the correlates of parental preferences. Section 6 concludes.

2. 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: mathematics, English language, sciences, social studies, and an essay. Students are allocated to

5. [Dobbie 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.

secondary schools by the MOE based on the SEA scores and the school preferences using the Deferred Acceptance mechanism summarized in Section 3.1 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% 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% 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, enrol 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 2 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 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,

6. There 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%).

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

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.

Labour market participation: We obtained the official registry of active contributors to the national retirement fund by 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% of students took the NCSE and 73.2% took at least one CSEC subject. The average student passed 3.2 CSEC subjects and 34.6% 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 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 vs. 88.3% 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 vs. 27.5% 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% qualify for tertiary education; while this is only accomplished by 11.6% of students at the least selective schools (column 5).

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

Moving to non-academic outcomes, 3.3% of the population had been arrested by age 18. Arrests are concentrated among males of which 5.8% had been arrested by age 18. Arrests rates are low (1.8%) among students from more selective schools, and are higher (4.7%) among students at the least selective schools. A similar pattern is observed for teen motherhood. While 6.9% of girls at the top schools had a live birth before age 19, as much as 15.2% of females at the bottom schools did. Finally, 75.5% 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.

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

TABLE 1
Summary statistics

	All schools	Males	Females	Above median	Below median
	(1)	(2)	(3)	(4)	(5)
Panel A: SEA data (cohorts: 1995–2012)					
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
Panel B: NCSE data (linked to SEA cohorts: 2006–2012)					
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
Panel C: CSEC data (linked to SEA cohorts: 1995–2011)					
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
Panel D: CAPE data (linked to SEA cohorts: 1999–2009)					
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
Panel E: Criminal records (linked to SEA cohorts: 1995–2010)—in percent					
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
Panel F: Birth records (linked to SEA cohorts: 2004–2010)—in percent					
Live birth by 19			10.11 (30.15)	6.86 (25.27)	15.19 (35.90)
Individuals			43,834	26,715	17,119
Panel G: Labour market data (linked to SEA cohorts: 1995–2006)—in percent					
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.

3. 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.

3.1. School assignments

The Ministry of Education (MOE) uses a Deferred Acceptance mechanism to create an initial set of school assignments for students. 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 20% 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.

3.2. Identification framework

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

$$Y_{ij} = v_j + \alpha_i + v_{ij} + \epsilon_{ij} \quad (1)$$

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).

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

9. See [Supplementary Appendix A](#) for a more detailed description of the school assignment process.

10. This mechanism generates higher cutoffs for schools that tend to be higher ranked by parents so that a more preferred school will be more selective. [Supplementary 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, [Supplementary 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 districts have access to both more and less selective schools.

11. Note 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.

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

where $\theta_{j,j'} \equiv [v_j - v_{j'}]$ reflects differences in value-added, $A_{j,j'} \equiv [\bar{\alpha}_{ij} - \bar{\alpha}_{i'j'}]$ 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 $E_{j,j'} \equiv [\bar{\epsilon}_{ij} - \bar{\epsilon}_{i'j'}]$ 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 $f(\text{SEA}_i)$, so that $E[A_{j,j'}|f(\text{SEA}_i), C=c] = 0$. However, because the assignment is conditional on $C=c$, $E[M_{j,j'}|f(\text{SEA}_i), C=c] = E[M_{j,j'}|C=c] = \mu_{j,j',c}$. 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. In such case, for each $c \in C$, in expectation, the difference in outcomes conditional on smooth functions of test scores would reflect true differences in value-added plus a differential match for individuals for whom $C=c$, as in (3).

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

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

Identifying assumptions: Under the framework above, $E[D_{j,j'}|f(\text{SEA}_i), C]$ 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 incoming SEA scores and (2) there are no differential average match effects (*this condition is satisfied under fixed value-added*). We show that these two conditions likely hold in Section 3.6.

3.3. 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 (Jackson, 2010; Cattaneo, Keele, Titiunik and Vazquez-Bare, 2021): (1) variation around the cutoffs for each school (based on applicants scoring just above and just below each cutoff) and (2) variation across cutoffs (based on *all* students including those far away from the cutoffs). We discuss each in turn.

3.3.1. 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

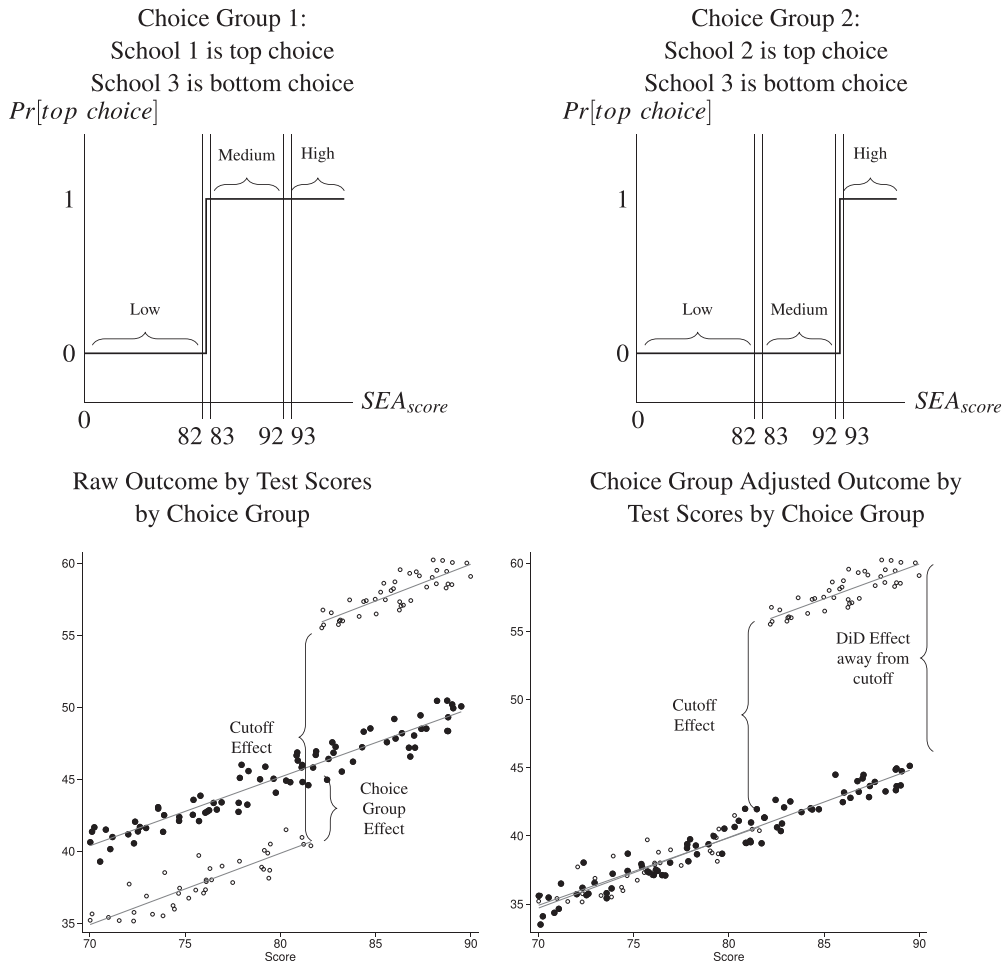


FIGURE 1
Exemplar of variation

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 labelled 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)—labelled 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 vs. 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).

(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 [Supplementary 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 one’s preferred school.

3.3.2. 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 cutoff 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

13. Intuitively, if students in different choice groups are assigned to the same school with similar incoming test scores, differences in outcomes across choice groups cannot be due 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 (1) the relationship between test scores and outcomes is the same across choice groups and (2) the difference in school effects is the same at all incoming test score levels, the relative value-added of all schools in the choice groups can be identified.

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 and the DiD models will yield the same relative school estimates (see Cattaneo *et al.* (2021) for a general discussion of this). In Section 3.6.1, we show that these assumptions likely hold.

3.3.3. 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*. 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). 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 3.6.3 we show that this additivity condition likely holds in our setting.¹⁵

3.4. Relying only on the identifying variation

We exploit variation both within and across cutoffs as discussed in Section 3.3. We refer to assigned school τ and attended school j . Based on the algorithm, students are tentatively assigned to school τ if they (1) had school τ in their list of choices, (2) scored above the cutoff for school τ , and (3) did not score above the cutoff for a more preferred school in their choices. Under the modelling assumptions discussed above (i.e. (1) scoring above the test score cutoffs is unrelated to potential outcomes, (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 by estimating (4) by ordinary least squares (OLS).

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

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 .

14. Suppose 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).

15. If 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.

$f(SEA_i)$ 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 ϵ_{itct} 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 cutoffs (using individuals away from the cutoff as outlined above). The estimated $\hat{\theta}_{\tau}^{ITT}$ s from (4) identify the average intention-to-treat (ITT) effect of being tentatively assigned to each school τ (relative to the same comparison school).¹⁸

3.5. 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.¹⁹ Identification of individual school effects requires one instrument per alternative (Kirkeboen, Leuven and Mogstad, 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.²⁰ We can obtain clean causal estimates for at least one outcome for 98.5% of all public secondary schools in the nation (i.e. 132 schools). The resulting two-stage least squares (2SLS) model is as follows:

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

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

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

17. In 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.

18. Note 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.

19. Noncompliance 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% of students tentatively assigned to their schools with any student of their choosing (see Supplementary Appendix A for a detailed description of this process). The second source of noncompliance is that students 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, this would render estimated impacts based on attended schools (without a convincing source of exogenous variation) biased.

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

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 that 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 labour 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.²¹ Using this same approach, we computed a "Low-Stakes Exams" index grouping both NCSE academic and non-academic performance. [Supplementary 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 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 labour market participation.

3.6. Testing the identifying and modelling assumptions

3.6.1. Testing identification assumption I: no selection to schools. As outlined in Section 3.2, 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 ([Supplementary Appendix A](#)). We now show that the no selection assumption also holds in our main models (which use variation both within and across cutoffs). As such, 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 would be obtained 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. We test this using a re-weighting method motivated by [Solon, Haider and Wooldridge \(2015\)](#) detailed in [Supplementary Appendix D](#). Because our DiD-type model uses variation among all admitted students, and schools admit students across a wide range of incoming scores, one can feasibly estimate relative school effects only among students with (or around) a particular incoming score. Accordingly, as suggested in [Solon et al. \(2015\)](#), we test for heterogeneous school effects

21. The first principal component represents the maximum variance direction in the data ([Jolliffe, 2002](#)). While the mean of standardized variables has been used in other studies (e.g. [Kling, Liebman and Katz, 2007](#)), there is no conceptual reason to put equal weight on each measure *in our context*. As it turns out, our measure and the mean of standardized variables has a correlation of 0.99 so that the distinction is minimal.

by incoming achievement by comparing school effect estimates weighted to be representative of impacts for those at the 75th, 50th, and 25th percentiles of the incoming test score distribution.²² In [Supplementary Appendix D](#), we show that relative school impacts are similar for these achievement levels.

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. 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. As we show in [Supplementary Appendix E](#), the effects are similar with and without such interactions.

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. As such, as detailed in [Supplementary Appendix F](#), 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*. Specifically, for each cutoff we estimate the change in actual outcomes right at the admission cutoff, and compare that with the change in the estimated value-added of the attended school right at the admission cutoff.²³ We show that the change in actual outcomes is largely the same as that predicted by our value-added estimates—validating the value-added estimates using the RD variation at the cutoffs.

3.6.2. Testing identification assumption II: no differential match effects. Now, we focus on the second identifying assumption of no differential match effects on average across schools. Specifically, following [Kirkeboen et al. \(2016\)](#), we find no evidence of what they call “*comparative advantage*” by showing that, on average, the effect of a School j relative to another School k is the same among those students who rank School j over k and among those who rank School k over j . The details of this test are shown in [Supplementary Appendix G](#).

3.6.3. Testing for additivity of school effects. Finally, as detailed in Section 3.3.3, even if we can identify relative school effects among individuals who make similar choices (as shown in Sections 3.6.1 and 3.6.2), our ability to compare *all* schools against each other relies on the assumption that school effects are additive. In [Supplementary Appendix H](#), we show that this condition holds empirically. That is, among pairs of schools that appear together in student choices, we estimate the relative effects of School m to that of School k (i.e. $\hat{\theta}_{m,k}^{ITT}$), and then using other pairs the effect of School m relative to other intermediate School l (i.e. $\hat{\theta}_{m,l}^{ITT}$) as well as the

22. 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.

23. A similar test was implemented in [Hastings, Neilson and Zimmerman \(2015\)](#) and [Beuermann and Jackson \(2022\)](#). This is also similar in spirit to the random assignment validation of school value added in [Deming et al. \(2014\)](#). See [Supplementary Appendix F](#) for further discussion of this test.

effect of School l relative to School k (i.e. $\hat{\theta}_{l,k}^{\text{ITT}}$). We then show that, on average, school effects are additive such that $\hat{\theta}_{m,k}^{\text{ITT}} = \hat{\theta}_{m,l}^{\text{ITT}} + \hat{\theta}_{l,k}^{\text{ITT}}$.

4. 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, Friedman and Rockoff, 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 ($\hat{\theta}_{j,\text{even}}^{\text{TOTIV}}$) and other including odd cohorts ($\hat{\theta}_{j,\text{odd}}^{\text{TOTIV}}$). Let $p \in \{\text{even}, \text{odd}\}$. 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 (7), we uncover maximum likelihood estimates of the variance of the persistent school impacts ($\sigma_{\theta^{\text{TOT}}}^2$) and of the transitory school impacts (σ_{μ}^2).²⁴

$$\begin{bmatrix} \theta_j^{\text{TOT}} \\ \mu_{jp} \end{bmatrix} \sim N \left(0, \begin{pmatrix} \sigma_{\theta^{\text{TOT}}}^2 I_J & 0 \\ 0 & \sigma_{\mu}^2 I_{(J \times p)} \end{pmatrix} \right) \quad (7)$$

Table 2 reports estimates of the standard deviation of the persistent school impacts for each outcome along with their 90% confidence intervals estimated by bootstrap (Column 1).²⁵ 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 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 5–7 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 those of attending high-impact Boston urban charter schools (Angrist, Pathak and Walters, 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 3 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

24. The 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.

25. To be conservative, we exclude outlier schools with estimates lying 4σ away from the median. In Supplementary Appendix Table B3, 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.

TABLE 2
Standard deviations and correlations of persistent school impacts

Outcome	School level ($\sigma_{\theta \text{TOT}}$)	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]
No live birth by 19	0.173 [0.147, 0.193]	−0.036 [−0.171, 0.087]	0.174 [0.021, 0.311]
Not arrested by 18	0.037 [0.032, 0.041]	0.282 [0.164, 0.419]	0.407 [0.210, 0.619]
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 $\text{weight}_i = (1 + \frac{(75 - \text{pct}_i)^2}{100})^{-1}$, 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).

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.²⁶

Teen motherhood: The standard deviation of the persistent school effects on teen motherhood is 0.173. The 90% 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 on teen motherhood.²⁷ Given that the teen live birth rate is around 10% on average, these represent large economically important relative impacts.

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%, this is a sizable reduction in teen arrests. Our estimates are larger

26. The 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.

27. In related work, Jackson (2019) finds that converting existing coeducation schools to single-sex reduced the teen birth rate by 4 percentage points. Dobbie and Fryer (2015) find that females admitted to a charter school in Harlem Children's zone are 10.1 percentage points less likely to be pregnant in their teens. Beuermann and Jackson (2022) find that attending a preferred school in Barbados decreases the teen motherhood rate by about 6 percentage points.

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 50%. The confidence interval does not include zero so that these school impacts are real and persistent over time.

Labour market participation: The final outcome we examine is participating in the formal labour market. We examine school effects on the likelihood that a student is observed with positive earnings in the formal labour 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 labour 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.

4.1. Is school quality unidimensional?

Many recent education policies (e.g. No Child Left Behind in the US or League Tables in the UK) 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](#)).²⁸ Following the notation in equation (7), 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 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 (8), one can estimate the correlation (net of estimation errors) between the effects on any two outcomes 1 and 2 (i.e. ρ_{12}) by maximum likelihood.

$$\begin{bmatrix} \theta_{1j}^{\text{TOT}} \\ \theta_{2j}^{\text{TOT}} \\ \mu_{1jp} \\ \mu_{2jp} \end{bmatrix} \sim N \left(0, \begin{pmatrix} \sigma_{\theta_1^{\text{TOT}} I_J}^2 & \rho_{12} \sigma_{\theta_1^{\text{TOT}} I_J} \sigma_{\theta_2^{\text{TOT}} I_J} & 0 & 0 \\ \rho_{12} \sigma_{\theta_1^{\text{TOT}} I_J} \sigma_{\theta_2^{\text{TOT}} I_J} & \sigma_{\theta_2^{\text{TOT}} I_J}^2 & 0 & 0 \\ 0 & 0 & \sigma_{\mu_1}^2 I_{(J \times p)} & 0 \\ 0 & 0 & 0 & \sigma_{\mu_2}^2 I_{(J \times p)} \end{pmatrix} \right) \quad (8)$$

28. The scatterplots of the raw school effects are shown in [Supplementary Appendix Figure B3](#).

Table 2 presents correlations between equally weighted (or average) school impacts on high stakes tests and the other outcomes (Column 2).²⁹ 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 (Supplementary Appendix Table B3). 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% (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).

One may worry, however, that the low correlations for school impacts across outcomes *may* reflect students who are marginal for different outcomes attending different schools. To assess this, we estimate school impacts weighted to be representative of students at the 75th percentile (where there is much common support as shown in Supplementary Appendix D).³⁰ By weighting the school impacts at the same point in the incoming 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 those around the 75th percentile of the incoming test score distribution*). We present the correlations between the estimated *weighted* school impacts on high-stakes and other outcomes in Table 2 (Column 3).³¹ 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. 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 (1) schools have economically meaningful impacts on a range of outcomes and (2) 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

29. Intuitively, 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 it 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 Supplementary Appendix Table B4).

30. That is, where pct_i is the percentile of student i in the SEA distribution, we estimate each school's treatment effect (θ_j^{TOTW}) while weighting each observation by $(1 + \frac{(75 - pct_i)^2}{100})^{-1}$. This puts heavy weight on students with incoming scores close to the 75th percentile and low weight on those far away from that percentile, yielding effects that are representative of those who receive more weight (Solon *et al.*, 2015)—i.e., those at the 75th percentile of the incoming test score distribution. See Supplementary Appendix D for the details of this procedure.

31. In Supplementary Appendix Table B3, we also show estimated correlations when we centre the weights at the 25th and 50th percentiles of the SEA distribution. Similarly, all estimates suggest moderate to low correlations.

school impacts on non-academic dimensions are economically meaningful and large. As such, the fact that parents do not choose schools that improve test scores (e.g. MacLeod and Urquiola, 2015; Abdulkadiroğlu *et al.*, 2020) may reflect parents choosing schools that improve other outcomes (that are weakly related to test score impacts). We explore these possibilities in Section 5.

5. ESTIMATING PREFERENCES FOR SCHOOLS

In this section, we will use the estimated school impacts and the set of secondary school choices to (1) examine the extent to which parents choose schools based on their causal impacts and (2) explore the extent to which they choose schools with causal impacts on outcomes other than high-stakes tests. As in all studies of this type (e.g. Hastings, Kane and Staiger, 2006; Avery, Glickman, Hoxby and Metrick, 2013; Burgess *et al.*, 2015; Abdulkadiroğlu *et al.*, 2020), 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 truthful choices 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.

5.1. 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 behaviour 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:

$$U_{ij} = U(X_i, Z_{ij}, \varepsilon_{ij}) = \delta(X_i, Z_{ij}) + \varepsilon_{ij}, \quad (9)$$

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 $U_{ij}^{r_{is}}$ indicate the utility parent i gets from school j that they ranked in position s ($r_i = s$), so that $U_{ij}^{r_{i1}}$ is their utility for the school ranked first, $U_{ij}^{r_{i2}}$ is their utility for the school ranked second, and so on. Let $U_{ij}^{r_{i0}}$ 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): $U_{ij}^{r_{ia}} > U_{ik}^{r_{ib}} \forall k \neq j \in J, a < b$ 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, 2022). 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 $U_{ij}^{r_{i1}} > U_{ik} \forall k \neq j \in J$ (Roth and Oliveira Sotomayor, 1992).³² 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 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.³³ 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 with 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

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

33. When 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).

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 (10) below.

$$\begin{aligned}
 Prob[r_{i1}, r_{i2}, \dots, R_i] = & Pr \left[\underbrace{(U_{ij}^{r_{i1}} | p_{ij} > U_{ik}^{r_{i0}} | p_{ik} \forall k \neq j \in J)}_{\text{Top choice preferred to all non-chosen schools conditional on admission probabilities}} \right. \\
 & \cap \underbrace{(U_{ij}^{r_{i1}} > U_{ik}^{r_{im}}, 1 < m, \forall m \in \{2, \dots, R_i\}) \cap \dots \cap (U_{ij}^{r_{iR_i-1}} > U_{ik}^{r_{iR_i}})}_{\text{Higher ranked chosen ex-post schools are preferred to lower-ranked chosen schools}} \left. \right]
 \end{aligned} \tag{10}$$

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).

5.2. Modified exploded multinomial logistic model

Equation (10) 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 (10) 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 a vector of deterministic components of school preferences, the *ex post* utility of student i from attending School j is

$$U_{ij} = \beta' Z_{ij} + \varepsilon_{ij} \tag{11}$$

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 these probabilities. 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 discuss these estimated probabilities in [Supplementary Appendix I](#). 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

$$U_{ij} | p_{ij} = \beta'_1 Z_{ij} + \beta'_2 (Z_{ij} q_{ij}) + \pi q_{ij} + \varepsilon_{ij}. \tag{12}$$

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 (11). 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 [McFadden, 1973](#); [Train, 2009](#)), and that of rational strategic behaviours, the probability that parent i submits a particular ranking over schools (i.e. Equation (10)) is simply a product of standard logit formulas.³⁴ Accordingly, where the parameter vector $\beta = [\beta_1, \beta_2, \pi]$, the log likelihood of observing all the choice lists for all parents is:

$$\log L(\beta) = \sum_{i=1}^N \log l_i(\beta) = \sum_{i=1}^N \log (\text{Prob}[r_{i1}, r_{i2}, \dots, R_i]). \quad (13)$$

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.³⁵ These two modifications to the conventional multinomial logit model allow us to anchor each individual’s choice set while also explicitly accounting for strategic behaviours.

5.3. 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

34. That 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:

$$\text{Prob}[r_{i1}, r_{i2}, \dots, R_i] = \frac{\exp(\beta'_1 Z_{ij}^{r_{i1}} + \beta'_2 (Z_{ij}^{r_{i1}} q_{ij}^{r_{i1}}) + \pi q_{ij}^{r_{i1}})}{\exp(\beta'_1 Z_{ij}^{r_{i1}} + \beta'_2 (Z_{ij}^{r_{i1}} q_{ij}^{r_{i1}}) + \pi q_{ij}^{r_{i1}}) + \sum_{k=1}^{J-R_i} \exp(\beta'_1 Z_{ik}^{r_{i0}} + \beta'_2 (Z_{ik}^{r_{i0}} p_{ij}^{r_{i0}}) + \pi p_{ik}^{r_{i0}})} \cdot \frac{\exp(\beta'_1 Z_{ij}^{r_{i1}})}{\sum_{k=1}^{R_i} \exp(\beta'_1 Z_{ik}^{r_{ik}})} \cdots \frac{\exp(\beta'_1 Z_{ij}^{r_{iR-1}})}{\exp(\beta'_1 Z_{ij}^{r_{iR-1}}) + \exp(\beta'_1 Z_{ik}^{r_{iR-1}})}.$$

35. Our model differs from [Hastings, Kane and Staiger \(2005\)](#) 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.

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 (12)) which represents the relationship between school preferences and choices when there are no strategic considerations—i.e., *when the probability of admission is one*.³⁶ 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; (1) an Impacts Only Model (which includes schools' causal impact estimates for all outcomes, peer quality, and log distance) and (2) 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 school rejection probabilities (but only when comparing the top choice to all unranked schools) and interactions of these probabilities 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.³⁷

To facilitate statistical inference, we pool estimates across different cells and report pooled coefficients and standard errors in Table 3.³⁸ Note that, for each model, we test for the significance of school impacts in six dimensions. To allow for a 5% type-I error among the six school impacts estimates, we use a simple Bonferroni 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 non-linear 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.

Our primary focus is to shed light on the relationship between choices and schools' impacts. However, our models include rejection probabilities, proximity, and peer quality, which are all potentially important determinants of choices. In [Supplementary Appendix L](#), we show that, *holding school impacts, peer quality, and proximity constant*, parents are more likely to choose schools with lower rejection probabilities—consistent with the rational strategic behaviours that we model. Also, in [Supplementary Appendix M](#), we show that choices are strongly related to proximity and peer quality—patterns that are consistent with existing studies (e.g. [Hastings et al., 2005](#); [Abdulkadiroğlu et al., 2020](#)). These 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.

5.3.1. Impacts on academic outcomes. Figures 2 and 3 plot the estimates separately for each outcome. In the top panels, we plot coefficients on schools' causal impacts from the Impacts

36. Our 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 [Supplementary Appendix J](#)).

37. See [Supplementary Appendix K](#) for the estimated coefficients on the school-level average outcomes from the Full Model.

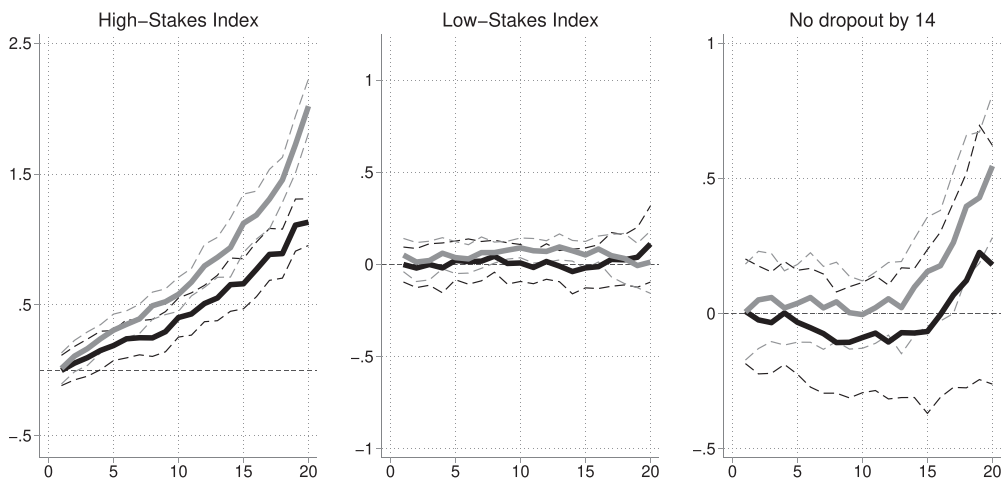
38. We report the mean of the coefficient across cells. The standard error of this mean is computed assuming independence across cells.

TABLE 3
Inference tests

	Impacts Only Model			Full Model			
	Pooled average (1)	Above median score (2)	Below median score (3)	Pooled average (4)	Above median score (5)	Below median score (6)	Female (7) Male (8)
Standardized value-added on							
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 (0.035)
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)
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)
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)
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)
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)

Notes: This table presents pooled averages of the choice model estimates on the standardized value-added of schools (i.e. β_1 from equation 12) 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.

(a) Causal Impact (Impacts Only Model)



(b) Causal Impact (Full Model)



FIGURE 2

Academic 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) \times (gender) cell, for two main models: (a) displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and (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.

Only Model. In the bottom panels, we plot coefficients on schools' causal impacts from the Full Model (that also includes school level averages).

We first discuss the relationship between school preferences and school impacts on high-stakes exams. Figure 2a 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 2b)—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 parents 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 (Table 3).³⁹ For those in the top 20% 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 km 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 5.4.

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).⁴⁰ Specifically, the Impacts Only Model (Figure 2a) 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 a t -statistic of 2.05 (Table 3), 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 (1) parents discerning school impacts on high-stakes exams but not on low-stakes exams or (2) parents not caring about school impacts on low-stakes tests precisely because they are low stakes.

39. Looking at the school average high-stakes scores ([Supplementary Appendix K](#)), 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 labour market for example.

40. An insignificant or small estimated coefficient could indicate that either parents do not value that particular school attribute or, alternatively, that parents care about it but they do not have enough information about it. We favour 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.

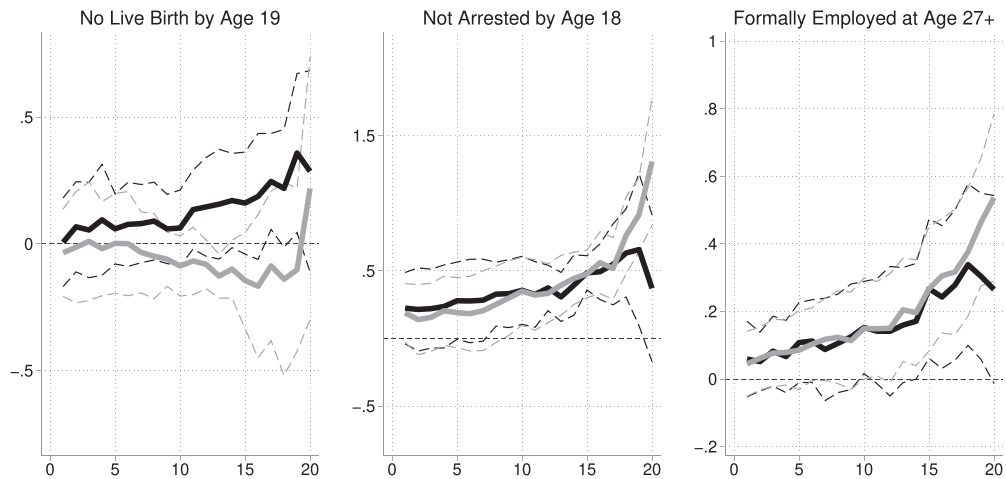
Our final academic outcome is dropout. As with low-stakes exams, the patterns suggest that parents are not systematically more likely to choose schools that causally reduce dropout. In the Impacts Only Model, the relationship is not different from zero for most cells, but there is some suggestive evidence that higher achieving parents may be more likely to choose schools that causally reduce dropout (Figure 2a). In the Full Model, for none of the gender-by-achievement cells can one reject that the coefficient on dropout impacts is zero (Figure 2b). 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 (Table 3). 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 could be particularly difficult for parents to observe and therefore respond to.

5.3.2. Impacts on non-academic outcomes. Next, we document the relationship between parent school choices and school impacts on non-academic outcomes in Figure 3. 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 (Table 3). 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 of 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% 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 (Table 3). This relationship is driven primarily by those below the median of the incoming achievement 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 females) 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 km to send their child to a school that was at the 85th percentile of the (non)-teen motherhood effect distribution vs. 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 of a male and female would be willing to increase their distance by about 11 and 4.5 km to send their child to a school that was at the 85th percentile of the (non)-teen motherhood effect distribution vs. one at the median, respectively.⁴¹ 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 births are correlated with some other school attribute that parents are responding to.

41. We 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 5–10 km 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 4, left panel).

(a) Causal Impact (Impacts Only Model)



(b) Causal Impact (Full Model)

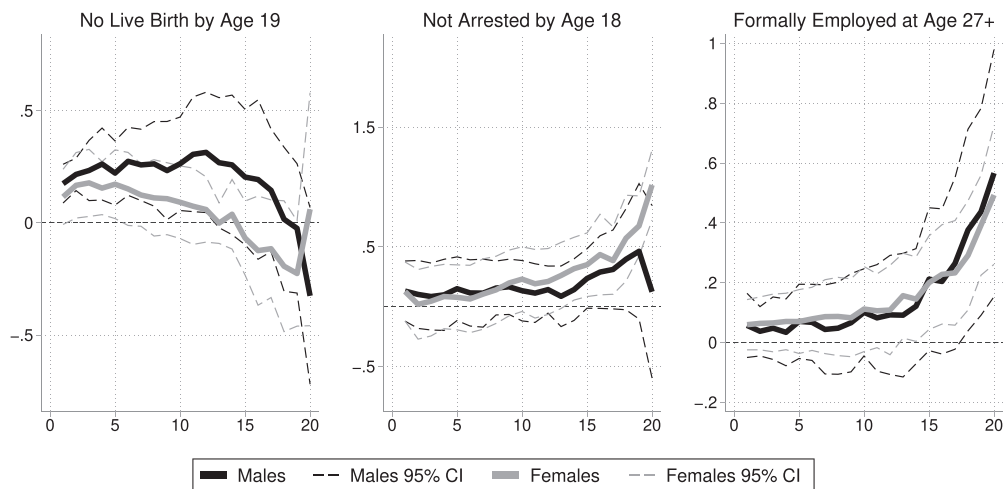


FIGURE 3

Non-academic 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) \times (gender) cell, for two main models: (a) displays estimates from the Impacts Only Model, which includes schools' causal impact estimates for all outcomes, peer quality, and log distance and (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.

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 3a. 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

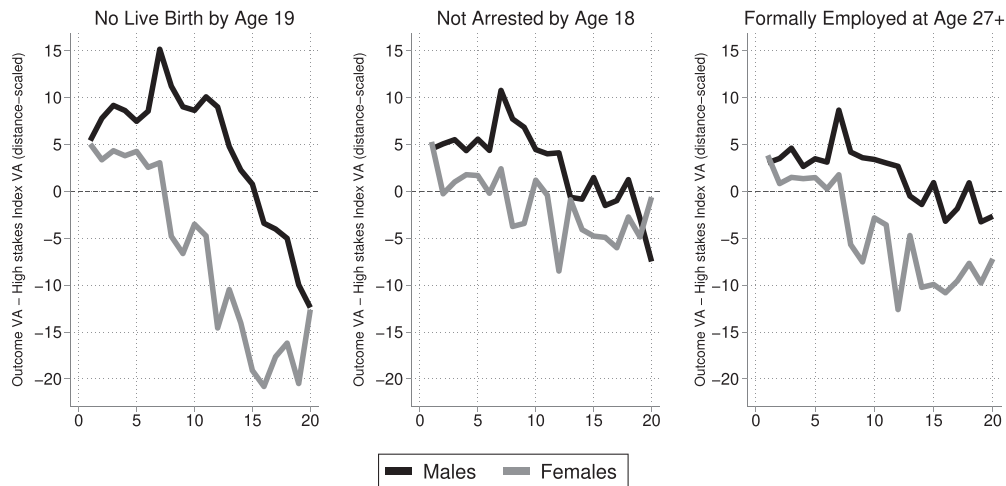


FIGURE 4

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) \times (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.

for most cells (the pooled t -statistic derived from Table 3 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 derived from Table 3 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 km to send their child to a school that was at the 85th percentile of the (non)-teen arrest impact distribution vs. one at the median, while that for those in the top decile is about 13 km. These results are consistent with parents preferring schools that reduce these behaviours (i.e. arrests) in their own children above and beyond peer achievement, proximity and average outcomes.

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 *victimization* 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 victimization. Second, we show that arrests do occur even among high achievers in [Supplementary Appendix Figure B4](#), such that being arrested is not irrelevant for any group. Third, we point out that being arrested is an equilibrium outcome so 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 gender) the patterns are consistent with one's intuitions. The middle panel of Figure 4 shows that low-achieving males may be willing to travel 5 km 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 arrests value added than girls, and parents of low-achieving students (particularly boys) appear to place more relative weight on arrests value added than those of high-achieving students—patterns that align well with one's 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 3). In the Impacts Only Model (Figure 3a), the pooled t -statistic (derived from Table 3) is 13.8—well above the multiple hypothesis adjusted threshold for rejecting no association. This is largely similar in the Full Model (Figure 3b)—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 km to send their child to a school that was at the 85th percentile of the employment impacts distribution vs. 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 with those for formal labour market participation, the right panel of Figure 4 suggests that low-achieving males and females may be willing to travel 4 and 1 km farther to attend a school with 1σ higher labour 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 km 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 labour market participation. Importantly, the correlations between school impacts on high stakes exams and impacts on arrests and formal labour 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.⁴² 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.⁴³ That is, we can rule out that our key results are driven by test score impacts having larger marginal effects for high-achieving children while non-academic impacts having larger effects on low-achieving children. Rather, our results likely reflect differences in preferences (or differences in information).⁴⁴

5.4. 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 motherhood, and labour 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 US (which offers over 700 programs at over 400 schools). Their setting is a context in which sub-optimal behaviours 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, 2022](#)). 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

42. Our 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 [Supplementary Appendix N](#)), 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.

43. In [Supplementary Appendix Table B5](#), 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.

44. 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.

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 schools' 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).

6. 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 labour 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 labour 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 schools' 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 schools' 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 non-academic outcomes.⁴⁵ The provision of such

45. In an experimental study in Romania, Ainsworth, Dehejia, Pop-Eleches and Urquiola (2020) find that distributing information on school academic value added led households (particularly those with low-achieving students) to attend more effective schools.

information may improve the decisions of all parents and could increase the *potential* allocative efficiencies and competitive benefits of school choice.

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 Boafio, 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 article benefitted from comments by Joshua Angrist, Samuel Berlinski, Matias Busso, Julian Cristia, Veronica Frisncho, Ofer Malamud, Norbert Schady, Diego Vera-Cossio, four anonymous referees, and seminar participants at the NBER, MIT, U. Chicago, LACEA, IADB, SALISES, Zurich, Lund, and Wharton.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online. And the replication packages are available at <https://dx.doi.org/10.5281/zenodo.6456606>.

Data Availability Statement

The data used in this article consist of confidential administrative data from different institutions in Trinidad and Tobago, which cannot be shared publicly. The data can be accessed by researchers by submitting a research proposal to each of them. Details on how to do this, as well as all replication scripts are available at DOI: <https://doi.org/10.5281/zenodo.6456606>.

REFERENCES

- ABDULKADIROĞLU, A., ANGRIST, J. and PATHAK, P. (2014), "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools", *Econometrica*, **82**, 137–196.
- ABDULKADIROĞLU, A., PATHAK, P., SCHELLENBERG, J. and WALTERS, C. (2020), "Do Parents Value School Effectiveness?", *American Economic Review*, **110**, 1502–1539.
- AGARWAL, N. and SOMAINI, P. (2018), "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism", *Econometrica*, **86**, 391–444.
- AINSWORTH, R., DEHEJIA, R., POP-ELECHES, C. and URQUIOLA, M. (2020), "Information, Preferences, and Household Demand for School Value Added" (NBER Working Paper 28267).
- ANGRIST, J., HULL, P., PATHAK, P. A. and WALTERS, C. R. (2020), "Simple and Credible Value-Added Estimation Using Centralized School Assignment" (NBER Working Paper 28241).
- ANGRIST, J., HULL, P., PATHAK, P. A. and WALTERS, C. (2021), "Credible School Value-Added with Undersubscribed School Lotteries" (NBER Conference Paper).
- ANGRIST, J. D., PATHAK, P. A. and WALTERS, C. R. (2013), "Explaining Charter School Effectiveness", *American Economic Journal: Applied Economics*, **5**, 1–27.
- ANGRIST, J. D., COHODES, S. R., DYNARSKI, S. M., PATHAK, P. A. and WALTERS, C. R. (2016), "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice", *Journal of Labor Economics*, **34**, 275–318.
- AVERY, C. N., GLICKMAN, M. E., HOXBY, C. M. and METRICK, A. (2013), "A Revealed Preference Ranking of U.S. Colleges and Universities", *The Quarterly Journal of Economics*, **128**, 425–467.
- BESEDEŠ, T., DECK, C., SARANGI, S. and SHOR, M. (2012), "Age Effects and Heuristics in Decision Making", *Review of Economics and Statistics*, **94**, 580–595.
- BEUERMANN, D. W. and JACKSON, C. K. (2022), "The Short and Long-Run Effects of Attending The Schools that Parents Prefer", *Journal of Human Resources*, **57**, 725–746.
- BOOKER, K., SASS, T. R., GILL, B. and ZIMMER, R. (2011), "The Effects of Charter High Schools on Educational Attainment", *Journal of Labor Economics*, **29**, 377–415.
- BURGESS, S., GREAVES, E., VIGNOLES, A. and WILSON, D. (2015), "What Parents Want: School Preferences and School Choice", *The Economic Journal*, **125**, 1262–1289.
- CARIBBEAN EXAMINATIONS COUNCIL (CXC). (1993–2016), "Caribbean Secondary Education Certification (CSEC), 1993–2016".
- CARIBBEAN EXAMINATIONS COUNCIL (CXC). (2005–2016), "Caribbean Advanced Proficiency Examination (CAPE), 2005–2016".

- CATTANEO, M. D., KEELE, L., TITIUNIK, R. and VAZQUEZ-BARE, G. (2021), "Extrapolating Treatment Effects in Multi-Cutoff Regression Discontinuity Designs", *Journal of the American Statistical Association*, **116**, 1941–1952.
- CHADE, H. and SMITH, L. (2006), "Simultaneous Search", *Econometrica*, **74**, 1293–1307.
- CHETTY, R., FRIEDMAN, J. N. and ROCKOFF, J. E. (2014), "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates", *American Economic Review*, **104**, 2593–2632.
- CHUBB, J. E. and MOE, T. M. (1990), *Politics, Markets, and America's Schools* (Washington, D.C.: The Brookings Institution).
- CRIME PROBLEM ANALYSIS BRANCH OF THE TRINIDAD AND TOBAGO POLICE SERVICE. (1990–2017), "Trinidad and Tobago's Arrest Records, 1990–2017". 10.5281/zenodo.6456606.
- DEMING, D. J. (2011), "Better Schools, Less Crime?", *The Quarterly Journal of Economics*, **126**, 2063–2115.
- DEMING, D. J. (2014), "Using School Choice Lotteries to Test Measures of School Effectiveness", *American Economic Review*, **104**, 406–411.
- DEMING, D. J., HASTINGS, J. S., KANE, T. J. and STAIGER, D. O. (2014), "School Choice, School Quality, and Postsecondary Attainment", *The American Economic Review*, **104**, 991–1013.
- DEUTSCH, J., GILL, B. and JOHNSON, M. (2020), "The Promotion Power Impacts of Louisiana High Schools (Executive Summary)", *Mathematica Policy Research Reports*, 1–44. <https://eric.ed.gov/?id=ED607741>.
- DOBBIE, W. and FRYER, R. (2020), "Charter Schools and Labor Market Outcomes", *Journal of Labor Economics*, **38**, 915–957.
- DOBBIE, W. and FRYER, R. G. (2015), "The Medium-Term Impacts of High-Achieving Charter Schools", *Journal of Political Economy*, **123**, 985–1037.
- FRIEDMAN, M. (1955), *The Role of Government in Education* (New Brunswick, N.J.: University of Chicago Press).
- GALE, D. and SHAPLEY, L. S. (1962), "College Admissions and the Stability of Marriage", *The American Mathematical Monthly*, **69**, 9–15.
- HAERINGER, G. and KLIJN, F. (2009), "Constrained School Choice", *Journal of Economic Theory*, **144**, 1921–1947.
- HANUSHEK, E. A. (1971), "Teacher Characteristics and Gains in Student Achievement: Estimation using Micro Data", *American Economic Review*, **61**, 280–288.
- HASTINGS, J., KANE, T. and STAIGER, D. (2005), "Parental Preferences and School Competition: Evidence from a Public School Choice Program" (NBER Working Paper 11805).
- HASTINGS, J., KANE, T. and STAIGER, D. (2006), "Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery" (NBER Working Paper 12145).
- HASTINGS, J., NEILSON, C. and ZIMMERMAN, S. (2015), "The Effects of Earnings Disclosure on College Enrollment Decisions" (NBER Working Paper 21300).
- HASTINGS, J. S., KANE, T. J. and STAIGER, D. O. (2009), "Heterogeneous Preferences and the Efficacy of Public School Choice" (Working Paper).
- HECKMAN, J. J., STIXRUD, J. and URZUA, S. (2006), "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*, **24**, 411–482.
- HOLMSTROM, B. and MILGROM, P. (1991) "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics and Organization*, **7**, 24–52.
- IYENGAR, S. S. and KAMENICA, E. (2010), "Choice Proliferation, Simplicity Seeking, and Asset Allocation", *Journal of Public Economics*, **94**, 530–539.
- JACKSON, C. K. (2010), "Do Students Benefit from Attending Better Schools? Evidence from Rule-Based Student Assignments in Trinidad and Tobago", *The Economic Journal*, **120**, 1399–1429.
- JACKSON, C. K. (2013), "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers", *Review of Economics and Statistics*, **95**, 1096–1116.
- JACKSON, C. K. (2018), "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes", *Journal of Political Economy*, **126**, 2072–2107.
- JACKSON, C. K. (2019), "Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood?", *Journal of Human Resources*, **56**, 1–39.
- JACKSON, C. K., PORTER, S., EASTON, J., BLANCHARD, A. and KIGUEL, S. (2020), "School Effects on Socio-emotional Development, School-Based Arrests, and Educational Attainment", *American Economic Review: Insights*, **2**, 491–508.
- JOLLIFFE, I. T. (2002), *Principal Component Analysis*, 2nd edn. (New York, NY: Springer).
- KANE, T. and STAIGER, D. (2008), "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation" (NBER Working Paper 14607).
- KAUTZ, T., HECKMAN, J. J., DIRIS, R., TER WEEL, B. and BORGHANS, L. (2017), "Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success" (NBER Working Paper 20749).
- KIRKEBOEN, L. J., LEUVEN, E. and MOGSTAD, M. (2016), "Field of Study, Earnings, and Self-Selection", *The Quarterly Journal of Economics*, **131**, 1057–1111.
- KLING, J. R., LIEBMAN, J. B. and KATZ, L. F. (2007), "Experimental Analysis of Neighborhood Effects", *Econometrica*, **75**, 83–119.
- MACLEOD, W. B. and URQUIOLA, M. (2015), "Reputation and School Competition", *American Economic Review*, **105**, 3471–3488.
- MACLEOD, W. B. and URQUIOLA, M. (2019), "Is Education Consumption or Investment? Implications for the Effect of School Competition", *Annual Review of Economics*, **11**, 563–589.

- MANSFIELD, R. K. (2015), "Teacher Quality and Student Inequality", *Journal of Labor Economics*, **33**, 751–788.
- MBITI, I., MURALIDHARAN, K., ROMERO, M., SCHIPPER, Y., MANDA, C. and RAJANI, R. (2019), "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania", *The Quarterly Journal of Economics*, **134**, 1627–1673.
- MCFADDEN, D. (1973), "Conditional Logit Analysis of Qualitative Choice", in Zarembka, P. (ed) *Frontiers in Econometrics* (New York: Academic Press) 105–142.
- MINISTRY OF EDUCATION. (1995–2012), "Secondary Entrance Exam (SEA)".
- MINISTRY OF EDUCATION. (2009–2015), "National Certificate of Secondary Education (NCSE)".
- MOUNTJOY, J. and HICKMAN, B. (2020), "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).
- NATIONAL INSURANCE BOARD OF TRINIDAD AND TOBAGO. (1980–2017), "Registrations in the National Insurance Board, 1980-2017".
- PATHAK, P. A. and SÖNMEZ, T. (2013), "School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation", *American Economic Review*, **103**, 80–106.
- PLACE, K. and GLEASON, P. (2019), "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).
- ROTH, A. E. and OLIVEIRA SOTOMAYOR, M. A. (1992), *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* (Cambridge: Cambridge University Press).
- ROTHSTEIN, J. M. (2006), "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions", *American Economic Review*, **96**, 1333–1350.
- SCHRAM, A. and SONNEMANS, J. (2011), "How Individuals Choose Health Insurance: An Experimental Analysis", *European Economic Review*, **55**, 799–819.
- SCHWARTZ, B. (2004), *The Paradox of Choice: Why More Is Less* (New York, NY: HarperCollins Publishers).
- SOLOMON, G., HAIDER, S. J. and WOOLDRIDGE, J. M. (2015), "What Are We Weighting For?", *Journal of Human Resources*, **50**, 301–316.
- SPEARMAN, C. (1904), "The Proof and Measurement of Association between Two Things", *The American Journal of Psychology*, **15**, 72–101.
- TRAIN, K. E. (2009), *Discrete Choice Methods with Simulation* (Cambridge: Cambridge University Press).
- TRINIDAD AND TOBAGO REGISTRAR GENERAL (2010-2016), "Trinidad and Tobago's Birth Records".