Northwestern | INSTITUTE FOR POLICY RESEARCH

Working Paper Series

WP-16-23

The Effect of Single-Sex Education on Academic Outcomes and Crime: Fresh Evidence from Low-Performing Schools in Trinidad and Tobago

Kirabo Jackson
Associate Professor of Human Development and Social Policy
IPR Fellow
Northwestern University

Version: May 2016

DRAFT
Please do not quote or distribute without permission.

ABSTRACT

In 2010, the Ministry of Education in Trinidad and Tobago converted 20 low-performing pilot secondary schools from coed to single-sex. I exploit these conversions to identify the causal effect of single-sex schooling holding other school inputs (such as teacher quality and leadership quality) constant. After also accounting for student selection, both boys and girls in single-sex cohorts at pilot schools score 0.14 higher in the academic subjects on national exams. There is no robust effect on non-academic subjects. Additionally, treated students are more likely to earn the secondary-school leaving credential, and the all-boys cohorts have fewer arrests. Survey evidence reveals that these single-sex effects reflect both direct gender peer effects due to interactions between classmates, and also indirect effects generated through changes in teacher behavior. Importantly, these benefits are achieved at zero financial cost.

The Effect of Single-Sex Education on Academic Outcomes and Crime: Fresh Evidence from Low-Performing Schools in Trinidad and Tobago

By C. KIRABO JACKSON[1]
4/27/16

In 2010, the Ministry of Education in Trinidad and Tobago converted 20 low-performing pilot secondary schools from coed to single-sex. I exploit these conversions to identify the causal effect of single-sex schooling holding other school inputs (such as teacher quality and leadership quality) constant. After also accounting for student selection, both boys and girls in single-sex cohorts at pilot schools score 0.14σ higher in the academic subjects on national exams. There is no robust effect on non-academic subjects. Additionally, treated students are more likely to earn the secondary-school leaving credential, and the all-boys cohorts have fewer arrests. Survey evidence reveals that these single-sex effects reflect both direct gender peer effects due to interactions between classmates, and also indirect effects generated through changes in teacher behavior. Importantly, these benefits are achieved at zero financial cost. (*JEL* I20, J00)

It is well-documented that boys and girls may react differently to the same educational interventions and environments.[2] Also, in some contexts, boys and girls may have better academic performance when exposed to more same-sex classmates (Whitmore 2005; Black et al 2013; Ooserbeek and van Ewijk 2014; Lu and Anderson 2015). These patterns, have led some to advocate for single-sex education – a form of tracking such that boys and girls are educated in separate classrooms or schools. While the merits of single-sex education have been hotly debated for decades, there remains a paucity of credible evidence on the effects of expanding single-sex education. To help fill this gap, this paper analyzes academic and crime effects of a policy experiment in Trinidad and Tobago under which twenty low-performing coeducational (coed) schools were converted to single-sex. This represents the first analysis of a large-scale policy to expand single-sex public education to low-achieving students.

In theory, holding other schooling attributes fixed, single-sex education may improve outcomes because (a) single-sex classrooms allow for instruction tailored to the specific needs of

[1] Jackson: Department of Education and Social Policy, Northwestern University, 2120 Campus Drive, Evanston 60208 (kirabo-jackson@northwestern.edu). This project was supported by the Spencer Foundation. I thank Brian Jacob, Julie Cullen, Gordon Dahl, Kitt Carpenter, Heyu Xiong, and Alexey Makarin for useful comments. I thank Carol Singh for invaluable project management and data collection efforts, and Igor Uzilevskiy, Patrick Peters, Diana Balitaan, Kevin Malis, Rodrigo Braga, Hao (Leo) Hu, Mathew Steinberg, Richard Yu, and Ben Henken for excellent research assistance. I am also grateful to Brenda Moore, Harilal Seecharan, and Peter Smith at TTMOE. All errors are mine.

[2] e.g. Jackson (2010) and Deming et. al. (2014) show that girls benefit more from attending good schools than boys.

each sex, (b) single-sex classrooms may allow teachers to better focus on instruction and (c) the presence of the opposite sex may be districting and affect social dynamics in ways not conducive to learning. If it were to work, single-sex instruction would be tremendously cost effective as it involves only the re-allocation of existing resources and no additional financial costs.[3] To help fix ideas, I present a model of single-sex schooling that nests these mechanisms above. The model highlights that single-sex schooling is neither always good nor always bad for either boys or girls. Moreover, the model suggests that the distinction between single-sex schools, single-sex cohorts, and single-sex classrooms may have to do with which mechanisms may be at play in each situation. However, to help make sense of the existing empirical literature, the model highlights the types of settings in which single-sex education will likely benefit or hurt students. I also derive testable implications from the model that can be evaluated empirically in my data.

Parents tend to perceive single-sex schools as being superior to coed schools (Jackson 2012, Park et al 2013), and more than one-third of Americans support having single-sex public schooling options (Howell, West, and Peterson 2008). These views are supported by observational evidence; In the US, single-sex charter schools (such as the well-known Urban Prep academies) that enroll low-income ethnic-minority students boast college-going rates well above average for schools serving similar populations (Chavous 2013). Also, graduates from all-women's colleges are more likely to pursue graduate degrees and have better labor market outcomes than those who attended similar coeducational institutions (Day 2008). Outside of the United States single-sex schools are also associated with better outcomes (Mael et al 2005). However, because of possible student selection to schools, it is unclear whether these apparent successes reflect causal relationships.

To address this selection problem, a literature has emerged that relies on quasi-random assignment of students to schools to compare outcomes at single-sex schools to those at coed schools among students who are similar in both observable and unobservable ways.[4] These studies have found positive effects of all-boys schools (Lee, et al. 2014), all-girls schools (Jackson 2012), or both (Park, Behrman and Choi 2013, Ku and Kwak 2013). However, because schools do not become single-sex at random, single-sex schools may differ from coed schools in unobserved ways that preclude a *like-with-like* comparison. Indeed, both Jackson (2012) and Ku and Kwak (2013)

[3] While the pecuniary costs to the school district is zero, there may be some welfare costs if parents or teachers do not like single sex education. However, evidence indicates that the opposite may be true.

[4]There is an older research literature in which researchers were unable to credibly disentangle the effects of single-sex schooling from the characteristics of the students who chose to attend single-sex schools (see Jackson 2012).

document that single-sex schools tend to have better observable characteristics than coed schools. As such, these cross-school comparisons (while informative about schools) are not informative about whether students benefit from single-sex education *per se*. Therefore, they do not speak to the policy-relevant question of how expanding single-sex education may affect student outcomes.

To address these limitations, I analyze a policy experiment in Trinidad and Tobago. In 2010, the Ministry of Education (MOE) identified 10 pairs (20 in total) of geographically close, similarly sized, low-performing coed public secondary schools. The aim was to improve the outcomes of low-performing boys at these schools. They selected one school in each pair to be converted to all-boys and the other to be converted to all-girls. The transition to single-sex schooling was gradual such that the incoming 6th grade cohorts after 2010 were single-sex while the previously admitted cohorts remained coed. The selected schools had no control over this decision. Also, to ensure a clean experiment, the MOE dictated that there be no other changes at these 20 schools. Because this experiment allows one to compare students who attended the same school under both coed and single-sex regimes, one can isolate the effect of adopting a single-sex policy from that of *unobserved* school-level differences that might exist between coed and single-sex schools.[5]

To analyze this policy experiment, I link student admission records prior to secondary school entry to national examination data taken in secondary school three years later, and the secondary school leaving examination taken five years later. These data allow me to analyze the effect of single-sex education on a broad array of academic outcomes. These outcomes include standardized test performance and course grades in a variety of subjects (including non-academic subtests such as physical education), the choice of which upper level subjects taken, high-school dropout and completing high school with a school leaving credential. I also link the secondary school admissions data with arrest records to present the first analysis of the effect of single-sex education on juvenile crime. I supplement these administrative data with survey data collected during the policy experiment to present evidence on mechanisms and test the theoretical model.

To identify the effect of the transition from coed to single-sex holding other school inputs

[5] In a related study, Ku and Kwak (2013) study outcomes at nine all-boys schools and four all-girls schools that changed from single-sex to coed and find that males performed better in coed environments while females fared worse. Unfortunately, why these schools converted to coed is unknown and data limitations do not allow the authors to rule out bias due to student selection or selective test taking. In related work, Booth, Cardona-Sosa and Nolen (2013) randomly assign students to all-girls and all-boys discussion sections in a University economics course. The authors find positive effects for females and no effect for males. It is unclear that these results will generalize to the context of secondary school children from disadvantaged backgrounds.

constant (i.e. a single-sex effect), I compare the outcomes of student cohorts who attended the same secondary school but who were admitted under coed versus single-sex regimes (i.e. before versus after the phased transition to single-sex). While this transition allows one to hold both other observed and other unobserved school inputs fixed, it does not ensure that the student populations are comparable across cohorts. To address this concern, I exploit discontinuities in the school assignment rules to isolate exogenous variation in school attendance and remove bias due to student selection to schools. I combine exogenous variation in school assignments with exogenous within-school changes in single-sex status to *compare the causal effect of attending an experimental school under the coed regime to the causal effect of attending that same school under the single-sex regime*. I present several empirical tests that this strategy is valid.

After accounting for both student selection and unobserved school-level differences, I find large positive effects of about 0.14σ on both boys' and girls' academic achievement on test taken three years after secondary school entry. Importantly, I show that the cross-school variation confounds any correlated school-level effects with the single-sex education effect –validating the use of within-school variation. While boys improve in both academic subjects and non-academic subjects, girls only improve in academic subjects. The positive effects for girls were similar across all academic subjects and were not just driven by math and science— counter to claims that single-sex education largely improves girls' outcomes in male-dominated subjects. The positive test score effects reflect improvements throughout the achievement distribution for boys and improvements in the lower tails for girls. Looking at non-test score outcomes, five years after secondary school entry, boys and girls are more likely to take advanced courses in the single-sex cohorts. Also, both boys and girls are more likely to earn a secondary-school-leaving credential, and boys are less likely to be arrested – suggesting that single-sex education improves both cognitive and softer skills (Jackson 2013, J. Heckman 1999). It is important to note that the test score effects are equivalent to reducing class size by about 20 percent or increasing teacher quality by 1.4 standard deviations. However, these benefit were gained at zero financial cost.

One limitation of the extant literature on single-sex education is a lack of evidence on underlying mechanisms. To help in this regard, I administered student surveys at pilot schools and a set of comparison schools in 2013, 2014, and 2015. Surveys reveal that gender peer effects are complex. For girls, surveys suggest positive direct peer effects in all-girls settings through less peer distraction, less peer disruption, and more peer learning. For boys, the direction of the direct peer effects is unclear; boys report feeling less socially anxious in all-boys setting, but also report learning less from their peers. Consistent with Lee et. al. (2014), there is also evidence of positive *indirect* effects in single-sex settings. There is limited evidence of greater alignment to each sex. However, there is evidence of efficiency gains to the more homogeneous single-sex classrooms because teachers spend more one-on-one time with students in single-sex classrooms despite no change in cohort size. The positive effects for both sexes echo Duflo, Dupas and Kremer (2011) who find that both low- and high-achievement students benefit from achievement tracking.

Note that the results presented speak to the effect of having all instruction in single-sex classes, which may be different from that of having a few single sex classes in a coed school. Also, I present the effect on schools of transitioning from coed to single-sex, holding teacher quality and other inputs fixed. This is a key parameter of interest. However, it may not represent of the long-run general equilibrium effects of introducing single-sex schooling to an education system if school inputs change in response to changes in the distribution of the student body in the long run.[6]

This paper makes a few new contributions. First, it lays out a model that outlines the contexts in which single-sex education is likely to be beneficial or harmful. Second, it is the first paper to identify the policy-relevant effect of converting coed schools to single-sex. Third, it presents evidence on mechanisms behind gender peer effects, distinguishing between direct and indirect effects. Fourth, it contributes to the tracking literature by analyzing the effects of gender tracking. Fifth, it adds to a small but growing literature documenting a causal link between education and crime. Finally, the findings speak to literatures on how certain types of schools can help economically disadvantaged youth (e.g. Fryer 2014, Deming, Hastings, and Staiger 2013).

The remainder of the paper is as follows: Section I lays out a simple theoretical model of single-sex schooling to inform and guide the subsequent empirical work. Section II describes the policy landscape, the policy experiment, and describes the data used in the study. Section III lays out the empirical strategy. Section IV presents the main empirical results, robustness checks and evidence on mechanisms. Section V presents a discussion and concludes.

I A Model of Single-Sex Education

The existing literature has laid out several explanations for how single-sex education may

[6] The data requirements for such an analysis are considerable. To address this question, one would need several education systems and one would have to exogenously introduced single sex schools into some of them and compare the outcomes of systems that had some conversions to those that did not. No such policy variation currently exists.

affect students. I present a model that nests these explanations and allows for three separate (not mutually exclusive) mechanisms through which single-sex education can affect student outcomes. The model shows that single-sex education can have different effects depending on context. Importantly, it highlights the conditions under which single-sex education effects may emerge.

Student Outcomes:

Student outcomes are given by (1) where Y_{ijg} is the outcome of student i with teacher j of gender $g \in \{girl, boy\}$, X_i is a set of incoming characteristics, G_i is an indicator equal to 1 if student i is a girl and 0 otherwise. $\bar{G}_j$is the share of girls in class with teacher j, the proportion of time that teacher j aligns instruction to girls is $p_j \in [0,1]$, and u_{ij} is a random mean zero error term.

(1) $$Y_{ijg} = \alpha G_i + X_i + f_g(\bar{G}_j) + h_g p_j + \psi\|2p_j - 1\| + u_{ij}.$$

There are three channels through which gender peer composition affects student outcomes. There is a direct gender peer effect and two indirect effects; the boutique effect and the focus effect. In (1), $f_g(\bar{G}_j)$ captures the **direct gender peer effect** that operates through peer interactions. Some studies find female classmates improve all students' outcomes— arguably because boys are disruptive (Hoxby 2000; Lavy and Schlosser 2011). However, others find that students have better outcomes when exposed to same gender peers— arguably because the opposite sex is distracting (Black et al 2013; Ooserbeek and van Ewijk 2014; Lu and Anderson 2015). To allow for differential responses to the proportion of female classmates, I include subscript g on f_g.

The term, $h_G p_j$ captures the idea that that certain teaching practices may benefit girls more than boys and *vice versa*. Where p_j is the proportion of time the teacher spends "aligning" classroom practices to the needs of girls, $h_{girls} \geq 0$ and $h_{boys} \leq 0$. The idea that students may benefit from similar peers because instruction can be aligned to their type undergirds the "Boutique" model of peer effects (Hoxby and Weingarth 2006), and is posited in Duflo, Kremer and Dupas (2011) model of ability tracking. This "**Boutique effect**" captures an indirect peer effect because the teacher's choice of p_j may depend on the gender composition in the classroom.[7]

The term $\psi\|2p_j - 1\|$ captures the idea that there may be some efficiency gains for teachers associated with focusing their instruction to one group (either boys or girls). Hoxby and Weingarth (2006) call this the "focus" mode of peer interactions. This "focus" term is motivated

[7] This gender alignment effect could be driven by the kinds of teaching examples used, the pedagogical practices employed, the discipline practices used, or even the ability level to which the class is pitched. As an example of gendered content, sports-based examples might be more engaging for boys than for girls.

by the finding that students tend to have better outcomes in more homogeneous classroom environments (e.g. Hoxby and Weingarth 2006, Ding and Lehrer 2007). The efficiency gains may be due teachers spending less time planning lessons for two separate audiences, or teachers spending less time managing the disengagement of those students who type is not being catered to at any given point in time. This "**focus effect**" term capture an *indirect* peer effect because the teacher's choice of alignment *p* may depend on the gender composition in the classroom.

Teacher's Choice of Alignment (p_j):

To understand how the *indirect* peer effects (i.e. the focus and boutique effects) operate across coed and single-sex settings, one must model how teachers chose gender alignment (p_j) as a function of the gender composition ($\bar{G}_j$). A teacher's payoff is an increasing function of the outcomes of her class $W(Y_{i\in j})$ so that $\partial W/\partial Y_i > 0 \;\; \forall i \in j$. Teachers chose how much time to spend aligning instruction to girls (p_j) in order to maximize their individual payoff.

<u>Proposition 1</u>: *Teachers will employ entirely male-aligned instruction in all-male classrooms and entirely female-aligned instruction in all-female classrooms.*[8] That is, p_j=1 in all-girls setting and p_j=0 in all-boys settings. The intuition is straightforward: Teachers prefer when their students have better outcomes. If there are no girls in the class, there is no benefit to aligning instruction to girls at all, so that teachers have an incentive to adopt the instructional style that improves the outcomes for all her students (who are all male). The analogous argument holds for all-girls classrooms.

<u>Proposition 2:</u> *In coed classrooms, instruction may be aligned to boys only, girls only, or some combination of the two.*[9] The choice of how to align instruction in mixed-gender classrooms depends on the parameters of the production function and the incentives faced by teachers.

Expected Benefits of Single-Sex Instruction:

Taking expectations of equation (1) for students in single-sex environments minus that for coed environments, will yield the single-sex treatment effect *β*, the expected difference in outcomes in single-sex environments (*s=1*) relative to coed environments (*s=0*). Because coed classrooms are roughly half female, I assume $(\bar{G}_j|s=0)=\frac{1}{2}$. The single-sex treatment effect for

---

[8] **Proof**: If the classroom is all-boys, then the for all $i \in j$, $h_g p_j < 0$, so that $(Y_{ij}|p_j = 0) > (Y_{ij}|p_j \neq 0)$. By assumption, $\partial W/\partial(Y_i) > 0 \; for\; all\; i \in j$, so that $W(Y_{i\in j}|p_j = 0) > W(Y_{i\in j}|p_j \neq 0)$. As such, teachers in all-boys classes will chose to align instruction to boys. Similarly, if the classroom is all-girls, then the for all $i \in j$, $h_g p_j > 0$, so that $(Y_{ij}|p_j = 0) < (Y_{ij}|p_j \neq 0)$. By assumption, $\partial W/\partial(Y_i) > 0 \; for\; all\; i \in j$, so that $W(Y_{i\in j}|p_j = 0) < W(Y_{i\in j}|p_j \neq 0)$. As such, teachers in all-girls classes will chose to align instruction to girls.

[9] See Appendix A for a discussion of this.

girls and boys is given by (2) and (3) below, respectively.

(2) $$\beta_{boys} = f_{boy}(0) - f_{boy}\left(\frac{1}{2}\right) + h_{boy}(1) - h_{boy}(p_j|s=0) + \psi - \psi\|2p_j - 1\|.$$

(3) $$\beta_{girls} = f_{girl}(1) - f_{girl}\left(\frac{1}{2}\right) + h_{girl}(1) - h_{girl}(p_j|s=0) + \psi - \psi\|2p_j - 1\|.$$

These equations show that the single-sex schooling effect for each sex depends on three factors.

The first factor is what I refer to as the "*direct peer interactions effect*" from $f_{boy}(0) - f_{boy}\left(\frac{1}{2}\right)$ and $f_{girl}(1) - f_{girl}\left(\frac{1}{2}\right)$. If all students benefit from more female classmates then, all else equal, girls will be better off in all-girls setting and boys will be worse off in all-boys settings. However, if both boys and girls benefit from more same-sex classmates, then both boys and girls in single-sex settings will enjoy positive direct peer effects.

The second factor is what I refer to as the "*boutique effect*" (i.e. the effect of having greater alignment to one's type) from $h_{boy}(1) - h_{boy}(p_j|s=0)$ and $h_{girl}(1) - h_{girl}(p_j|s=0)$. Because teachers will align instruction to girls in all-girls setting and boys in all-boy's settings, the magnitude of the "boutique" effect depends on the extent of alignment to girls or boys in coeducational settings. The fact that the single-sex effect includes $(p_j|s=0)$ is because boys will only enjoy an alignment effect if coed classes are girl aligned and *vice versa*. This is important, because the choice of alignment in coed settings depends on teachers' incentives in coed settings which may vary from context to context. If teachers split their attention in coed settings (i.e. $0 < p_j < 1$), both boys and girls in single-sex settings may benefit from the boutique effect.

The third factor is what I refer to as the "*focus effect*" (i.e. the positive effect of having a teacher focus her instruction to only one type) and is summarized with $\psi - \psi\|2p_j - 1\|$. This effect is maximized in single-sex settings, but the magnitude in coed setting is unclear. The more teachers split their time between aligning instruction to both girls and boys in coed settings, the greater is the benefit to single-sex schooling. Importantly, the focus model effect could lead to larger benefits to single-sex education for both boys and girls, but depends on the level of alignment in coeducational settings. If teachers were already fully aligned to any one group in coed settings, then there would be no additional focus effect in the single-sex settings.

In sum, the model underscores that single-sex schooling is neither always good nor always bad and depends on the mechanisms at play in the specific context. This is important for thinking about how single-sex schools may differ from single-sex classrooms within coed schools. If teachers have a greater incentive to align instruction to one sex in single-sex schools than in single-sex classrooms within coed schools, one may see larger benefits to single-sex schools than single-sex classrooms. Also, if the direct gender peer interactions in the classroom are affected by the gender composition of the school, single-sex classrooms may have different effects from that of single-sex schools. While the model does not make strong predictions about what one may observe in any one situation, it does organize thinking around the effects, and shows that the single-sex effect depends on three key factors; (a) the size and direction of the direct gender peer effects (b) the change in alignment of instruction between coed and single-sex settings, and (c) the size of the teacher "focus effect". Even though I do not have independent variation in each of these causal mechanisms, I employ survey data to present evidence on these causal pathways.

II. The Trinidad and Tobago Context and the Pilot Program

The Trinidad and Tobago education system evolved from the English system. Secondary school begins in form 1 (grade 6) and ends at form 5 (grade 10). The main island, Trinidad, is where all the pilot schools are located and is roughly 50 miles long and 37 miles wide. All of the pilot schools are government (traditional public) schools.[10] These schools provide instruction from forms 1 through 5 and teach the national curriculum. There are two externally graded exams that students take in secondary school. These are key outcomes in this study. The first is the National Certificate of Secondary Education (NCSE) taken at the end of form 3 (grade 8) by all students (both in public and private school) in eight subjects.[11] The second exam is the Caribbean Secondary Education Certification (CSEC) taken at the end of form 5 (grade 10).[12] These are equivalent to

[10] There are two types of public secondary schools: Government schools and Government Assisted schools (Assisted schools). Government schools are fully funded and operated by the Government while Assisted school are run by private bodies (usually a religious board) and *at least* half of their expenses are paid for by the Government. Along all other dimensions, Government and Government assisted schools are identical. Assisted schools are similar to charter schools in the United States. Unlike in many countries where private schools are often of higher perceived quality, private schools in Trinidad and Tobago account for a small share of student enrollment and tend to serve those who "fall through the cracks" in the public system (roughly 2 percent). Students who attend private secondary schools have test scores a third of a standard deviation lower than the average SEA student, and half a standard deviation lower than the average among those students who take the CSEC exams.

[11] Students are assessed using regular continuous assessments. The students' final score is a weighted average of course grades in that subject (submitted by their teachers) and an externally graded exam. The MOE awards certificates to all students who attain a combined final mark of at least 60% for each of the eight core subjects.

[12] There is a third exam taken after the CSEC. This is the Caribbean Advanced Proficiency Examinations CAPE taken at the end of upper 6th form (12th grade). This is the equivalent of the British A-levels and passing scores can count for college credit at many U.S. Universities. This is considered a post-secondary credential.

the British Ordinary levels exams.[13] Students seeking to continue their education take five or more subjects, and all testers take the English language and math exams.[14]

A useful feature of Trinidad and Tobago education system is that students are assigned to secondary schools by the MOE. At the end of primary school (after grade 5, typically at age 11) students take the Secondary Entrance Assessment (SEA) examinations, and each student lists four ordered secondary school choices. These choices and their SEA score are used by the MOE to assign them to schools using a deferred acceptance algorithm (detailed in Section III). This assignment algorithm creates test-score cut-offs for each school that determines admission to that school. The variation in school attendance driven by these cut-offs is outside the control of students or parents and plays a key role in isolating the school effects from selection effects in this study.

II.A. Description of the Data

This project uses the administrative SEA data from 2007 through 2012. These data include scores on the national examination taken at the end of grade 5, the school choices made by the student prior to sitting the SEA exam, and the administrative secondary school assignment. The data also include age, gender, primary school, and religious affiliation. The final dataset contains information on 119,279 students across the seven SEA cohorts. The key outcomes under study are test scores in grade 8. As such, the SEA data are linked to the NCSE data for 2009 through 2015 by full name and date of birth.[15] The NCSE data contain scores earned on the NCSE exams taken at the end of form 3 (grade 8) when the typical student is 15 years old. The NCSE data contain scores earned in Mathematics, English, Spanish, Science, Social Studies, Arts, Physical Education, and Technology. This allows for an analysis of effects on a variety of subjects beyond mathematics and language (as in most studies) that cover both academic and non-academic content.

This paper also studies effects on non-test score outcomes. Because all students are required by law to take the NCSE exams (at both public and private schools), the NCSE data include information for students who are missing scores but are registered as full time students.

[13] There are 31 CSEC subjects covering a range of academic subjects such as Physics, Chemistry and Geography, and vocationally related subjects such as Technical Drawing and Principles of Business and Office Procedures.

[14] The CSEC exams are accepted as an entry qualification for higher education in Canada, the UK and the United States. After taking the CSEC, students may continue to take the Caribbean Advanced Proficiency Examinations (CAPE), at the end of grade 12, which is considered tertiary level education but is a prerequisite for admission to the University of the West Indies (the largest University in the Caribbean). The CAPE is the Caribbean equivalent of the English Advanced Levels (A-Levels) examinations.

[15] The match rate of 88 percent is consistent with the national dropout rate of 20 percent by age 17.

Accordingly, one can infer dropout using these data where dropout is defined as not being officially enrolled in any school (public or private) in the nation. To present suggestive evidence on high school completion, the SEA data are also linked to the administrative examination data for the 2012 through 2015 CSEC (taken 5 years after secondary school entry) by name and date of birth.[16] Because the first treated cohort entered school in 2010 and is expected to take the CSEC in 2015, this allows for a suggestive analysis of the CSEC only for the first treated cohort. The SEA data are also linked by name and date of birth to official arrest records. These data cover all arrests between January 1st 2000 and February 1st 2015. Roughly 2 percent of males in the SEA sample during the pre-transition years had been arrested by the age of 16, and more than 10 percent of those who dropped out of secondary school had ever been arrested-- so that this is a meaningful outcome. Summary statistics for the analytic dataset are provided in Table 1.

Columns 1 and 4 show the means and standard deviations of the main variables for boys and girls separately. Roughly half of the sample is male (50.6 percent). On average, about 5.6 percent of both male and female SEA takers is a repeat taker. One notable pattern is that females tend to outperform males on average. Females score one third of standard deviation higher than males on the SEA exam at the end of 5th grade. This gender gap continues through high school such that girls score about one-third of standard deviation higher than males on the NCSE math exam at the end of 8th grade. The gender gap is larger for English, Spanish, and social studies where females score about half a standard deviation above males and smallest for physical education where females outperform males by 0.089 standard deviations. These patterns indicate that females outperform males on average in all subjects, but that the relative male underperformance is greater in language and humanities subjects (which are typically female dominated fields). As in other nations, the high school dropout rate is higher for males than females: 0.145 for females compared to 0.2 for males. These high school completion numbers are comparable to those in the United States; the high-school non-completion rate in the United States was 20 percent in 2012. The gender differences in secondary school completion are even larger. While 42 percent of females earn the secondary school leaving certificate required for tertiary education entry, only 28 percent of males do. As expected, the likelihood of being arrested is much higher for males than females. Roughly 1.1 percent of males are arrested by the age of 16 while

[16] The match rate of 82 percent is lower than the match rate for the NCSE (taken two years earlier) and is in line with the purported national dropout rate of 20 percent by age 16.

only 0.2 percent of females are. These figures paint a picture of male academic underperformance, male discipline problems, and moderate dropout rates for both sexes.

II.B. The Pilot Program

The MOE was very concerned about the underperformance of males. However, the MOE noticed that the boys who attended elite single sex schools (studied in Jackson 2012) tended to have good outcomes, and thus decided to experiment with single-sex education. In March of 2010, the MOE identified 20 low-performing co-educational government schools to be converted to single-sex by September 2010 (there were 90 government schools in the country at the time). These pilot schools were selected in pairs and were in close proximity to each other. One school in each pair was chosen to be converted to all-girls and the other all-boys. The decision of which schools were chosen was done entirely by the Ministry without the approval or consultation of the schools.[17] These 20 pilot schools were to be converted on a phased basis (with each successive incoming cohort) to single-sex schools over a period of five years. Commencing with the 2010 SEA placement process, incoming 6th grade students were assigned to single-sex classes in the pilot schools. To avoid disrupting the existing students, students who were previously admitted to coed cohorts remained in coed cohorts, while those admitted after the transition were admitted to single-sex cohorts. In 2010, the 6th graders were in single-sex classes, while grades 7 through 10 were in coed classes. In 2011, the new incoming 6th graders were single-sex, the initial treated cohort who were now in 7th grade continued to be single-sex, while those earlier cohorts (who were now in grades 8 through 10) continued to be in coed classes. This phased cohort-level transition was to continue until all pilot schools were single-sex in all grades.[18]

The schools for the pilot were chosen based on three criteria: (1) The pilot schools would be non-selective schools, (2) the pilot schools had to be traditional government schools (as opposed to government assisted schools—i.e. charter schools), (3) the pilot schools had to be close to another pilot school of similar selectivity, and (4) because the pilot pair of schools would need to

[17] There has been criticism of this initiative, particularly from the Trinidad and Tobago Unified Teachers Association (TTUTA). It has been put forward that there was not prior consultation and there is insufficient time for proper preparation for what will be a major change in this large number of schools.

[18] The status of each grade in pilot school by year is presented in Appendix B. However, in 2013 there was a change in government, there was growing frustration from the teachers' union that they were never consulted on the transition, and single-sex pilot was abandoned. In 2013, the Ministry of Education announced that the incoming 6th graders in 2014 would be coed. As before, the composition of admitted cohorts were preserved so that those who were admitted into the single-sex regime remained in single-sex cohorts even after the policy was abandoned. By 2014, the 6th grade cohort was coed while all the older grades were single-sex.

take in half of the students who would have attended to pair, pilot schools had to be close to another pilot school of similar size.[19] To demonstrate that the selection of pilot schools followed the stated MOE criteria, Figure 1 plots the likelihood of being a pilot school for different groups of schools by distance to the nearest government school. As shown in Figure 1, only traditional government schools were chosen for the pilot. Among traditional government schools, only those of below average selectivity were chosen for the pilot, and among the non-selective traditional government schools the likelihood of being a pilot is strongly associated with being close to another government school. In fact, among non-selective traditional government schools, over 40 percent of those that were within one kilometer of another government school were chosen as pilot sites, and none more than 2.5 kilometers from a government school was chosen. In a regression, school type, school selectivity, and distance "explains" over half of the variation in pilot school status.[20] In sum, pilot school status was involuntary. Because the MOE selected pilot schools based on known criteria, one can be confident that pilot schools were not cherry picked or chosen based on a trajectory of improving or worsening outcomes. However, to assuage any lingering concerns, I show that pilot schools did not exhibit any differential pre-trending in outcomes in Section IV.

Table 1 presents summary statistics for boys and girls at the pilot schools before and after the transition. As expected, boys and girls who attended pilot schools prior to transition had weaker incoming preparedness than the average student. Specifically, boys and girls at pilot schools pre-transition scored 0.12 and 0.59 standard deviations lower on the SEA than the average male and female, respectively. Interestingly, after the transition, there is evidence of negative selection to all-boys schools and positive selection to all-girls schools. Boys at all-boys pilot schools after the

---

[19] To show that this is consistent with the actual pilot schools chosen by the MOE, I obtained address data for each secondary school in Trinidad and Tobago, geocoded each address, and computed the shortest distance (as the crow flies) between each secondary school and the closes government school. I then merged this data in with the average incoming SEA scores of student assigned to each school. Addresses were obtained from a variety of sources: contacting individual schools for address data, websites for individual schools, and official databases with school information. Addresses were geocoded using the reported longitude and latitudes from google maps. In a few cases street addresses could not be geocoded so that area centroids and street mid points were used instead. Accordingly, there may be some small errors in distance calculations.

[20] Even though school type, school selectivity, and distance can explain much of the variation in Pilot school status one may wonder why some were chosen and other were not. One reason is that the distance calculation used are based on ass the crow flies. The MOE would have used distance measures based on travel time. This would introduce some measurement error that would reduce the explained variation. A second reason is that some schools that were close to a government school and were non-selective were a different size from the closest school – making the transition less feasible. Third, some schools are clustered such that at times three schools are closer to each other but only two schools (one pair) could be chosen for the pilot. Finally, because the MOE wished to have pilot schools in different areas where several schools are clustered closer to each other only two (one pair) would be chosen for the pilot.

transition scored 0.123 standard deviations lower than boys at the same schools prior to the transition. In contrast, after the transition, incoming SEA scores increased by 0.156 standard deviations among girls at pilot all-girls schools. The patterns show evidence of changing selection within schools such that a comparison of mean outcomes before versus after the transition will be biased. I discuss my approach to credibly identifying the single-sex schooling effect in Section III.

III. Empirical Framework

The transition policy caused pilot schools (*T*=1) to go from coed in the pre period (*Post*=0) to single-sex in the post period (*Post*=1). Comparison schools (*T*=0) remained coed in both periods. I lay out an empirical framework that defines the parameter of interest, describes how it can be identified in the current context, and discusses how it relates to existing estimates in the literature.

I model outcome Y_{ij} of student *i* at school *j* as below.

[4] $$Y_{ij} = (T_j \times Post)\sigma + \theta_j + \beta X_{ij} + \pi \times Post + \varepsilon_{ij}$$

In [4], T_j is equal to 1 if the student attends a pilot school and zero otherwise, *Post* is equal to 1 in the post transition period and zero otherwise, θ_j, is a school fixed-effect, X_{ij} is a matrix of student level covariates, and ε_{ij} is a random error term. The parameter of interest, σ, is the marginal effect of attending a pilot school in the post transition period relative to the pre-transition period. This is the effect of introducing single-sex education, net of student selection effects, and holding all school inputs (e.g. teacher quality) fixed – i.e. the causal effect of single-sex education. This is the policy-relevant parameter that captures the change in school effectiveness caused by the adoption of single-sex instruction. As I explain below, existing studies that rely on cross-school comparisons fail to identify to parameter of interest and may confound σ with student and school factors.

The typical approach to uncovering σ has been to rely on variation across schools to identify the effect of attending a single-sex school (with the hopes that this identifies the single-sex schooling effect). A naïve approach of this sort would rely on the across-school variation and compare the outcomes of students at single-sex schools (*T*=1, *Post*=1) to those at coed schools (*T*=0, *Post*=1). To see what this naïve comparison identifies, consider the conditional expectation of the difference in outcomes between observationally similar students who are in single-sex and coed environments in the post period *because they attend different schools* in [5] below.

[5] $$E(\hat{\sigma}_{Naive}) \equiv E[Y_{ij}|T=1, Post=1] - E[Y_{ij}|T=0, Post=1] \equiv \sigma + E\big[\Delta\theta_j\big] + E[\Delta\varepsilon_{ij}] \ .$$

This naive estimate reflects the true single-sex effect, σ , plus unobserved school-level differences that exist between single-sex (i.e. pilot) and coed (i.e. non-pilot) schools $E[\Delta\theta_j] \equiv E[\theta_j|T=1] - E[\theta_j|T=0]$, plus a selection term $E[\Delta\varepsilon_{ij}] \equiv E[\varepsilon_{ij}|T=1, Post=1] - E[\varepsilon_{ij}|T=0, Post=1]$. Because (a) students are not randomly assign to schools, and (b) schools do not become single-sex at random it is clear that $E(\hat{\sigma}_{Naive}) \neq \sigma$.

A recent literature has relied on natural experiments to remove bias due to student selection to single-sex schools (e.g. Jackson 2012, Park et al 2013, Lee et al 2016) in the cross-section. These studies improve upon the naïve estimates and rely on exogenous variation across schools such that the selection term is equal to zero in expectation (i.e. $E\left[\Delta\varepsilon_{ij}\right] = 0$). These studies compare outcomes of students who are similar in both observed and unobserved dimensions who are in single-sex and coed environments because they attend different schools. Using only the post transition data, this cross-school selection free estimate, $\hat{\sigma}_{selection,post}$, is such that

$$[6] \qquad E\left(\hat{\sigma}_{selection,post}\right) \equiv E[Y_{ij}|S=1, Post=1] - E[Y_{ij}|S=0, Post=1] \equiv \sigma + E[\Delta\theta_j] .$$

As [6] makes clear, these selection free cross-school estimates *do capture the causal effect of attending a single-sex school relative to a coed school*. However, it also makes clear that such estimates do not identify the parameter of interest, σ, unless $E\left[\Delta\theta_j\right] = 0$. That is, if there are unobserved school-level differences across single-sex and coed schools, these studies, while they convincingly address the selection problem, will not identify the policy-relevant parameter of interest, σ, the causal effect of single-sex education *all else equal*.

To address this concern, I exploit the fact that the pilot program allows one to observe the same schools under both single-sex and coed regimes at different points in time (or for different cohorts).[21] If one can exploit exogenous variation across the same schools both pre and post transition (where some schools change single-sex status over time), one can remove *both* the selection bias and also bias due to unobserved school-level attributes. This is my approach.

Consider now, using the same exogenous variation used to estimate $\hat{\sigma}_{selection,post}$, (the selection free difference between the pilot schools and the non-pilot schools), but estimating this same cross-school difference in the pre-transition period (when all schools are coed). As before, the selection term is equal to zero, and there are unobserved school-level differences that exist

[21] Intuitively, one could compare post-transition students in single-sex classrooms to pre-transition coed students within the same school –holding school level differences constant. Sadly, while such an approach will remove school-level differences, this strategy does not address the selection problem.
$E[Y_{ij}|T=1, Post=1] - E[Y_{ij}|T=1, Post=0] = \sigma + E[\varepsilon_{ij}|T=1, Post=1] - E[\varepsilon_{ij}|T=1, Post=0] \neq \sigma$.

between pilot and non-pilot schools $E[\Delta\theta_j]$. However, now there is no single-sex schooling effect. Specifically, the resulting selection-free cross-school estimate in the pre-period is

[7] $$E(\hat{\sigma}_{selection,pre}) \equiv E[Y_{ij}|S=1, Post=0] - E[Y_{ij}|S=0, Post=0] \equiv E[\Delta\theta_j].$$

The selection free estimate in the pre period captures the difference in school characteristics between pilot and non-pilot schools ($E[\Delta\theta_j]$). The selection-free estimate in the post period captures the single-sex schooling effect plus the difference in school-level characteristics between pilot and non-pilot schools ($\sigma + E[\Delta\theta_j]$). The selection-free pilot-school effect estimate in the post period ($\hat{\sigma}_{selection,post}$) minus the selection free pilot school effect estimate in the post period ($\hat{\sigma}_{selection,pre}$) identifies the single-sex schooling effect (σ). As discussed in Jackson (2010, 2012, 2014), there is an assignment mechanism that is used to assign students to schools each year. I detail how this mechanism can be used to remove selection in both pre and post transition cohorts to facilitate the difference-in-difference type estimate outlined above.

III.A. Student Assignment to Secondary Schools

At the end of primary school (grade 5), all students take the Secondary Entrance Assessment (SEA) exams (typically in March). Before sitting the examination, students submit a list of four ordered school secondary school choices. The tests are sent to the Ministry of Education (MOE) where they are externally scored. The test score and the school choices are used by the MOE to assign students to secondary schools using a deferred acceptance algorithm (Gayle and Shapley 1962).[22] This algorithm creates a test score cut-off for each government school below which applicants are not admitted and above which applicants are admitted (Jackson 2010).

In practice, however, there is not *full* compliance with the cutoffs because the MOE makes administrative assignments in certain circumstances. For example, students who score below the cutoffs for all the schools in their choices receive an MOE assignment that may not comply with the cutoffs. Also, when schools are forced to restrict capacity for unforeseen reasons (e.g. one school was closed down for a year due to fire and another could not open due to flooding), students receive administrative assignments to accommodate these circumstances. It is important to stress that the noncompliance in school assignments is due to MOE adjustments that are outside the control of parents and students. As such, one can use the cut-offs to estimate a selection-free effect of attending a pilot school using a fuzzy-regression discontinuity type design.

[22] For full details on how students are assigned to schools see Appendix C.

The cut-offs for each school are not reported by the MOE. However, because I have access to the administrative assignment (which is outside the control of the students or their parents), there *are* real cut-offs that were used to assign students, and any non-compliance with the cut-offs are orthogonal to students' attributes, the cut-offs can be recovered empirically. For each of the 20 pilot schools, I regress whether an applicant is *assigned* to the school *by the MOE* on an indicator variable denoting scoring above each possible test score.[23] I select the cut-off that yields the largest *F*-statistic to be the cut-off for that school. If one used the actual schools attended, one might worry that this empirical approach would result in endogenous cut-offs. However, this approach uses the MOE assignment and *not the actual school attended* to infer the location of the cut-offs. In most cases there is sufficient compliance that there is little question about the location of the cut-offs.

To show that this procedure closely approximates where the actual cut-offs likely were, the top panel of Figure 2 shows the likelihood of receiving the administrative assignment to pilot school X (the school's name is confidential) as a function of applicants' incoming SEA scores in 2009 (the year before the transition) and 2010 (the year after the transition). The estimated cutoff is indicated by the dashed grey vertical line. It is clear that in this school the administrative assignments follow the cut-off rules virtually all the time. As such, it is clear where the cut-off was for this school in both years. However, the MOE did not assign all students to schools using the cutoff rules in all years. To show this, the lower panel presents the likelihood of attending a different pilot school in 2008 and 2011. Even though the MOE assignments were not in full compliance with the cutoffs for this school in these specific years, within a relatively small range of values, the location of the cutoffs is easily inferred from the data (somewhere around the 52rd and 56th percentile in 2008 and around the 53rd percentile in 2011. Consistent with this, the regression that imposes the cut-off at the 53th percentile yields the highest *t*-statistic in both years, indicating that this score is the most likely to have been use by the MOE given the distribution of administrative assignments. While some schools' assignments are better approximated by the cut-off rule than others (as shown in Figure 2), all results are robust to only using cut-offs that explain more than seventy percent of the variation in the administrative assignments.

III.B. Identification Strategy

The main strategy to remove selection to schools within each cohort is to use the cut-offs

[23] Note that an applicant to a school is a student that lists that school in their choices and was not already assigned to a school higher up in their choices.

described above in a fuzzy regression discontinuity design. As shown in in Figure 1, among students who chose a pilot school in a given year, the likelihood of being assigned to a pilot school increases in a discontinuous manner as one's score goes from below to above the cut-off for that pilot school. If the locations of the cut-offs are orthogonal to student characteristics, and the effect of test scores on outcomes are smooth through the cut-offs, one can attribute any sudden jumps in outcomes through the cut-offs to the sudden increased likelihood of being assigned to, and therefore attending, a pilot school in a given year.

To isolate the effect of transitioning from coed to single-sex from other school attributes, I can (a) use the cut-offs in pre-transition years to obtain a regression discontinuity (RD) estimate the causal effect of attending a pilot school prior to the transition ($\hat{\sigma}_{selection,pre}$), (b) use the cut-offs in post-transition years to obtain a regression discontinuity (RD) estimate of the causal effect of attending a pilot school *after* the transition ($\hat{\sigma}_{selection,post}$), and then (c) use these two estimates compute the *change* in relative effectiveness of pilot schools after the transition – i.e. the single-sex schooling effect ($\hat{\sigma}_{selection,post}$-$\hat{\sigma}_{selection,pre} = \sigma$). The proposed estimator of the single-sex schooling effects is a difference-in-regression-discontinuity (DiRD) design.

The validity of the DiRD approach requires that both the RD identifying assumptions be satisfied (i.e. smoothness of potential outcomes through the cut-off), and the DiD identifying assumptions be satisfied (i.e. common trends between treatment and comparison schools and no other changes at pilot schools). However, the DiRD design has the additional assumption that, in expectation, the Local Average Treatment Effect (LATE) of the pilot school fixed effect is the same before and after transition.[24] Because the transition may have caused some students who would have applied to pilot schools to no longer do so, the marginal students (i.e. students close to the cutoff) may be different prior to the transition than after. If the characteristics that differ between the marginal students before and after the transition are also those are systematically associated with having larger or smaller pilot school effects, then one could not simply "difference out" the school effect between the two RD estimates and the DiRD estimates would be biased. In section IV.1 I present evidence that each of these sets of assumption is likely satisfied.

[24] Note that this does not require that there is no heterogeneity in the pilot school effect or that the marginal students are the same over time. This requires the much weaker condition that the change in the marginal students over time is unrelated to the underlying benefits to attending pilot schools. The condition one worries about it where those close to the cut-offs after the transition derive much larger benefits from attending pilot schools than those who were at the cutoffs before the transition. I show that this is not likely the case. If anything, the opposite may be true.

Within a regression setting this DiRD estimate can be obtained by the two stage least squares (2SLS) model outlined in equations [8], [9], and [10]. Y_{ijc} is the outcome of person i in school j in SEA cohort c, and $Above_{ij}$ is an indicator equal to 1 if a student applied to pilot school j and also scored above the cut-off for pilot school j. π_{choice} is a fixed effect for fixed effect for the student choices. This choice fixed effect denotes the unique combination and order of the four school choices so that all comparisons are made among students who chose the exact same set of schools in the exact same order.[25] This is important because the assignments based on the cut-offs are exogenous *conditional on school choices.* All SEA cohorts admitted to pilot schools after 2010 (i.e. Post=1) are single-sex. All variables are defined as in [4].

[8] $$Y_{ijc} = (\widehat{T_J \times Post_\iota})\sigma + \hat{T}_j + \beta X_i + \pi_{choice} + \pi_c + \varepsilon_{ijc}$$

[9] $$\hat{T}_j = \Sigma(Above_{ic} \times I_{y=c})\rho_{c,1,1} + \rho_{2,1}X_i + \pi_{choice,1} + \pi_{c1} + \varepsilon_{ijc,1}$$

[10] $$(\widehat{T_J \times Post_\iota}) = \Sigma(Above_{ic} \times I_{y=c})\rho_{c,2,1} + \rho_{2,2}X_i + \pi_{choice,2} + \pi_{c2} + \varepsilon_{ijc,2}$$

The endogenous regressors are attending a pilot school, T_j , and attending a pilot school after the transition, $T_j \times Post$. To remove selection bias, I instrument for these two endogenous regressors with scoring above the cut-off for a pilot school interacted with the SEA cohort of the students, $\Sigma(Above_{ic} \times I_{y=c})$. Conditional on school choices, scoring above the cut-off for a pilot school in the SEA cohorts prior to 2010 will isolate exogenous variation in pilot school attendance during the pre-transition years (when all schools were coed), and scoring above the cut-off for a pilot school in SEA cohorts in 2010 and after will isolate exogenous variation in pilot school attendance during the post-transition years (when pilot schools were single-sex). Because it is important model the smoothness of outcomes through the test score cut-offs, the incoming student characteristics X_{ij} include a fourth-order polynomial of incoming SEA scores for each SEA cohort. Because the model includes the estimated effect of attending a pilot school across all years, the coefficient on $(\widehat{T_J \times Post_\iota})$ identifies the parameter of interest σ – the causal effect of the change in the effectiveness of pilot schools after the transition to single-sex. Standard errors are adjusted for clustering at the assigned school level.

III.C. Illustrating the DiRD Variation Visually

The DiRD estimate is the difference between fuzzy-RD estimates obtained before and after

[25] For example students who chose schools A,B,C, and D are in a different group from those who list B,A,C, and D.

transition. Figure 3 presents the first stage for all the pilot schools combined.[26] The figure shows the likelihood of *attending* a pilot school during the pre-transition years on the left and the post transition years on the right, as a function of applicants' scores relative to the cut-off. The figure also includes the 95% confidence interval of a model that fits the data with a fourth order polynomial of the relative score and an "above-cut off" indicator. The coefficient on the "above-cut off" indicator is presented along with the *p*-value associated with the null hypothesis that there is no shift through the cut-off. While the data are somewhat noisy, for both the pre and the post years, there is a visible shift in the likelihood of attending a pilot school through the cut off (a relative score of 0). The increase in likelihood of attending a pilot school is 50 and 71 percentage points in the pre and post-transition periods, respectively (both *t*-statistics are well above 10).

To illustrate the mechanics of the DiRD estimate, Figure 4 shows the analogous figure for the change in 8th grade math scores through the same cut-offs. The left figure show that math scores are 0.01σ lower through the cut-off for the pre-transition period (p-value=0.64), so that scoring above the cut-off for a pilot school in the post-period is associated with a 50 percentage point increase is pilot school attendance and 0.01σ *lower* math scores. Using only this discontinuity variation, the Wald estimate of the pilot school effect is 0.01/0.75 = 0.02σ in the pre period. The right figure shows that math scores are 0.1σ higher through the cut-off for the post-transition period (*p*-value=0.01), so that scoring above the cut-off for a pilot school in the post-period is associated with a 71 percentage point increase is pilot school attendance and 0.1σ *higher* math scores. Using only this discontinuity variation, the Wald estimate of the pilot school effect is 0.1/0.71 = 0.14σ in the post period. The DiRD estimate is the difference between the RD effect in the post period (which includes the single-sex effect and school level differences) minus that for the pre period (which includes only school level differences) which is 0.16σ (i.e. 0.14- + 0.02). In words, the causal effect of attending a pilot school relative to a comparison school on math test scores increased by 0.16σ after the transition to single-sex relative to before.

This example illustrates the logic of, and presents a visual representation of, the DiRD variation used in the 2SLS model outlined above. As long as (a) the cut-off variation removes selection bias (the RD identifying assumption), (b) there were no other school-level changes at

[26] Following Pop-Eleches and Urqiula (2013) and Jackson (2010), I show multiple cut-offs in a single graph by taking the applicants for each pilot school in each year (140 cut-offs; 20 schools over 7 years), re-centering their SEA scores around the cut-off for that school in that year and then stacking the data for each cut-off.

pilot schools during the transition (the Difference-in-difference identifying assumption), and (c) the Local Average Treatment Effect (LATE) of attending a pilot school is the same for those at the cut-off before and after the transition, this variation will yield the causal effect of single-sex schooling on outcomes. I show that this variation is likely valid in Section IV.

IV. Main Results

Table 2 presents the main regression results for both boys and girls simultaneously. Point estimates are presented for attending an all-boys pilot school, an all-boys pilot school after the transition, an all-girls pilot school, and attending an all-girls pilot school after the transition. The coefficient on "Pilot(boys)*Post" is the estimated *change* in the effectiveness of pilot schools after the transition to all-boys, and the coefficient on "Pilot(Girls)*Post" is the estimated *change* in the effectiveness of pilot girls' schools after the transition to all-girls.

I discuss the effect for the boys and girls in turn. Column 1 shows the raw difference in 8th grade scores (with no controls) for all-boys pilot schools both pre and post transition. The coefficient on attending a pilot all-boys school is -0.462 (p-value<0.01) and attending a pilot all boys school after transition (relative to pre transition) is -0.287 (p-value<0.01). Because boys tend to score below girls on average, part of this might simply be a gender composition effect. To asses this possibility Column 2 presents the same raw comparisons only for boys. Models that only compare boys with boys reveal a similar pattern of results, but now the naïve transition estimate falls by half to -0.14 (p-value<0.01). That is, after the transition boys scored 0.14σ lower on the math NCSE exams than before. To a casual observer, outcomes for boys at schools that transitioned to all-boys were demonstrably worse after the transition than before.

Column 3 presents results that control for students' incoming test scores. The estimated transition effect goes from negative and statistically significant to positive – indicative of negative selection to the all-boys pilot schools after transition. One would expect individual SEA scores to remove a large amount of self-selection bias. Despite this, OLS without choices may still be biased because students may know more about their ability and aspirations beyond their SEA scores, which may be noisy. Adding student choices should remove this additional bias. Column 4 presents results that condition on both incoming test scores and student choices. Adding school choices increases the estimated positive transition effect to 0.126 (*p*-value<0.05). This rich model that addresses selection on observables, shows that boys at pilot schools after the transition score

0.126σ higher than observationally similar boys who at pilot schools before the transition.

To address bias due to selection in unobserved dimensions, Column 5 presents the 2SLS models that isolate exogenous variation in school attendance due only to a students' score relative to that of the test score cut-offs of pilot schools. The first stage F-statistics for all endogenous regressors are well above 20 so that there is no weak identification problem. These 2SLS model yields a transition effect of 0.158 (p-value<0.05). This is similar to the implied single-sex effect calculated in the illustrative visual example presented in Section III.C. The fact that the 2SLS estimate is larger than the OLS estimates is consistent with the general pattern that as one accounts for more sources of selection, the estimated transition effect becomes more positive for the all-boys schools. That is, as one adopts more and more credible identification strategies, the estimated all-boys effect is more positive. While the naïve comparisons might be suggestive of negative all-boys effects, the results show that this is driven by strong negative selection to all-boys pilot schools after the transition in both observed and unobserved dimensions.

One key innovation of this paper is to account for school-level characteristics by exploiting within-school variation in single-sex status over time. To show the importance of this, Column 6 shows the estimated single-sex schooling effect for the post transition cohorts only (i.e. using only the cross-school variation). The estimated single-sex school effect for boys is negative (-0.053) and is not statistically significantly different from zero. This demonstrates how relying on cross-school variation (even if it is clean) confounds single-sex status with other school-level differences, and may be *very* misleading about the causal effect of single-sex education.

Column 1 also shows the raw difference in 8th grade scores for all-girls pilot schools both pre and post transition. The coefficient on attending a pilot all-girls school is -0.525 (p-value<0.01) and attending a pilot all-girls school after transition (relative to pre transition) is 0.292 (p-value<0.01). This shows that, similar to all-boys pilot schools, the all-girls pilot schools had substantially worse outcomes on average than other schools (i.e. pilot school were lower performing schools). However, unlike the raw naïve transition effect for all-boys pilot schools, performance at these school was better after the transition to all girls. Models that only compare girls with girls reveal a similar pattern of results, but now the naïve transition estimate falls to 0.137 (p-value<0.01). Outcomes for girls at schools that transitioned to all-girls were better after the transition than before. However, this may not reflect a causal effect.

Adding student test scores prior to secondary school entry (Column 3) has no economically meaningful effect on the estimates transition effect. The model that conditions on both incoming test scores and student choices (Column 4), shows that girls who attended pilot schools after the transition score 0.13σ higher than observationally similar girls at pilot schools before the transition. These effects are surprisingly similar to those for all-boys schools. To address bias due to selection in *unobserved* dimensions, Column 5 presents the 2SLS models that isolate exogenous variation in school attendance due only to the test score cutoffs. The 2SLS model yields a transition effect of 0.119 (p-value<0.05). This is similar to the estimated all-boys effect. Both naïve comparisons and more sophisticated models indicate positive all-girls effects on test scores. While the magnitude of the effects varies across models, the positive all-girls effects is a robust finding.

Because one cannot reject the null hypothesis that the all-girls transition effects is the same as the all-boys transition (p-value=0.62), Column 7 presents the 2SLS transition effect for boys and girls combined. The coefficient is 0.143 (p-value<0.01), so that on average, boys and girls who attended pilot schools after the transition scored 0.143σ higher than similar girls and boys at pilot schools before the transition. Under the identifying assumptions, this relationship is causal.

IV.A. Effects on Other Subjects

To illustrate the nature of the variation, the analysis thus far has focused on math scores. However, students take several subjects for NCSE certification. Table 3 presents the estimated 2SLS single-sex effect in academic subjects (Math, English, Spanish, Science, Social Studies) and nonacademic subjects (Arts, Physical Education, Technology). Looking at the combined single-sex transition effect (top panel), there are positive and statistically significant effects in Math, English, Spanish, Science, and Social studies. Across these subjects, the estimated effects are similar and lie between 0.09σ and 0.17σ. The effects on the nonacademic subjects are not statistically significant. However, the point estimates for these subjects are positive. The results indicate positive single-sex effects on academic subjects and suggestive evidence of positive effects on the non-academic subjects. To summarize the effects on academic subjects, I computed the mean of the academic scores. The point estimates in specification 9 show that the transition increased boys' and girls' test scores in academic subjects by 0.141σ (p-value<0.01). The lower panel (models 11 through 20) explores the effects for girls and boys separately. By and large, there are no systematic differences by gender. However, the point estimates suggest that the benefits in academic subjects were larger for boys, and that boys also improved in non-academic subjects.

IV.B. Specification Checks

As discussed in Section III, the DiRD variation requires that both the RD identifying assumption is valid and the DiD identifying assumption is valid. I summarize several tests of these assumptions here (see Appendix D for a full discussion of these tests). In order for the RD design to be valid, there should be no gaming of the cut-off and outcomes must vary smoothly through the cut-offs. One diagnostic for gaming is to test for non-smoothness in density through the cut-off. Such tests reveal no change in density through the cut-offs. To test for smoothness in outcomes through cutoffs, one can test for smoothness of latent outcomes (as proxied by covariates) through cutoffs. In such tests, I find that predicted outcomes (i.e. fitted values from a regression of math scores on repeater status in 5^{th} grade, student religion, primary school district, selectivity of the student's primary school, and month of birth) vary smoothly through the cut-offs. Another concern is that the effects could have happened by random chance by choosing any arbitrary cut-offs. To test for this, I created placebo cutoffs for each school (1000 times) and found that less than three percent of the placebo cut-offs generated effects on academic test scores as large as the actual cut-offs. In sum, all of the specification checks indicate that the RD variation is clean.

Even if the RD variation is clean, for the DiRD estimates to represent the causal effect of single-sex education requires that there were no changes in the pilot schools over time that coincided with the pilot program. As discussed previously, confounding changes within schools is unlikely because (a) the individual schools had no control over when they would become pilot schools, (b) the government stipulated that no other changes take place in these schools, and (c) schools were not made aware of the changes until the summer preceding the change so that schools had no time to react before the start of the school year. Even so, one may worry that the pilot schools were already on an upward trajectory prior to transition. To show evidence that this does not drive the results, Figure 5 presents the RD estimates for attending a pilot school for each year between 2006 and 2012 on academic scores. Each data point is an RD point estimate, and the 90 percent confidence interval is presented for each RD estimate. The RD estimates for 2006 through 2009 hover around -0.12, and there is no indication of an upward trend at pilot schools prior to the transition. Consistent with the positive single-sex effects for boys and girls, the pilot school effects in 2010, 2011 and 2012 are higher than that in the pre-treatment years and hover around 0.02. The difference between the RD estimates pre and post transition suggest that the transition increased the RD pilot school effect by roughly $0.12+0.02=0.14\sigma$—in line with Tables 2. Figure 7 presents visual evidence that the DiD identifying assumption is likely valid.

Even though the MOE clearly stipulated that there be no changes, as a final check on the DiD assumptions of no other changes, I used survey data an all a sample of teachers at all schools obtained in 2009 and 2013 to see if any systematic personnel changes occurred during the transition. In a simple difference in difference regression model (Appendix Table D1), between 2009 and 2013 the transition schools saw no differential change in the percentage of female teachers, or the percentage of teachers with a Bachelor's degree. There is a marginally statistically significant effect on year of birth at the all-male pilot schools. This is consistent with sampling variability and is in the opposite direction of what would be required to generate a positive transition effect for the all-male pilot schools. I also explore if differences in class size (proxied by cohort size) can explain the results. In a simple model predicting initial cohort size with school fixed effects and year fixed effects, the post transition pilots have no statistically significant difference in cohort size (and the effect for the all-male and all-female pilot school are in opposite directions). The results corroborate the notion that there were no other changes at pilot schools.

Even if the RD identifying assumptions and the DiD assumptions are both satisfied, in order for the DiRD strategy to uncover the transition effects requires that the Local Average Treatment Effects (LATE) identified in the pre-period and the post period for pilot schools be the same. Specifically, if the students who were at the cut-offs after the transition benefited more from attending pilot schools than those before, due to treatment heterogeneity it would appear that the transition increased the effectiveness of pilot schools over time.[27] To test for this possibility empirically, I estimate models that interact attending a pilot school with various observed student characteristics. These include incoming test scores, the number of single sex school in the student's choices, the selectivity of schools in the student choices, the selectivity of their primary school, whether the primary school was single sex, and predicted math scores (based on repeater status in 5th grade, the student's religion, the primary school district, selectivity of the student's primary school, and month of birth). Figure 6 present the estimated single sex transition effect in models that allow for treatment heterogeneity in these dimensions (models 1 through 8). Across all models

[27] Prima facie evidence that this does not drive the results is in Figure 5. The RD estimate is very similar for the 2010, 2011 and 2012 cohorts. While the 2011 and 2012 cohorts made their school choices with full knowledge of the transition, the 2010 cohort did not. Students in the 2010 cohort who were assigned to a pilot school were automatically assigned to the sister/brother school in each pair that matched their sex. As such the applicant to the pilot schools in 2009 and 2010 will be very similar. Note, however, that the MOE allowed parents of children assigned to pilot schools in 2010 to opt out of their MOE assignment in 2010. This does not invalidate the RD method, but it does leads to lower compliance with the cutoffs in 2010.

the singe sex schooling effect is similar to the main models. In fact, in the only models that are different, the estimated single-sex effect is larger than in the main specification.

Finally, one may worry that the positive effects of the transition may be driven by schools that are listed just below the pilot schools being much worse after the transition to single sex than before (i.e. the comparison school are worse after the transition). I test for this in two ways. First, I computed the percentile rank of all schools in incoming SEA scores and tested for whether the percentile ranking of the school just below the pilot school (among pilot school applicants) were worse after the transition. On average among pilot school applicants, the next ranked school is 0.85 percentile points (p-value<0.05) lower ranked than before the transition. Looking among individuals within 20 percentile points of the cut-off for a pilot school (i.e. the population used for identification), the next ranked school is 0.6 percentile points lower after the transition than before. This effect is not economically significant and is not statistically significant at the 10 percent level. Even though the comparison schools are not much worse after the transition, it is helpful to see that the estimates are robust to accounting for the selectivity of the next ranked school explicitly. To do this I estimate a model that interacts pilot school attendance with the selectivity of the next ranked school (Models 9 and 10 in Figure 6). Importantly, the estimated single-sex transition effect is unchanged in such models. Given the unusually rich set of covariates used (including revealed preferences for single sex schools), all these results in Figure 6 suggest that any heterogeneity in the estimate LATE pilot school effect over time does not drive the main results.

IV.C. Distributional Effects

The results thus far have focused on average effects. However, one may wonder if these improved outcomes are driven by improvements among the abler students at the upper end of the achievement distribution, or among the weaker students at the bottom. To test for this, I put each student's score into a test score quartiles, and I estimate the main 2SLS models where the main outcomes is scoring at each quartile. I present the coefficient estimates for boys and girls in Figure 7 along with the 90 percent confidence interval for each regression estimate.

The math gains for boys appear to be driven by improvement across the achievement distribution. After the transition, boys were 6 percentage points less likely to be in the bottom quartile, and 6 percentage points more likely to be in the top quartile. The patterns in English for boys suggest larger improvements at the top of the achievement distribution; Boys 8 percentage points more likely to be in the top quartile. In contrast, the benefits for girls are concentrated at the bottom of the achievement distribution for both subjects. After the transition, girls were 3.6 percentage points less likely to be in the bottom quartile in Math and were about 8 percentage points less likely to be in the bottom quartile in English. In both subjects, girls who would have scored in the bottom 25 percent scored in the middle of the distribution. Overall, the results suggest that single-sex education improved outcomes throughout the achievement distribution for boys with more pronounced improvements among the least able students for girls.

IV.D. Other Outcomes

Looking beyond achievement, I analyze dropping out of secondary school by the end of 8^{th} grade and being arrested. The 2SLS regression estimates are presented in Table 4. The top panel present results for boys and girls combined. Column 1 shows that there is no effect of the transition on dropout. To examine if there is some effect heterogeneity, the lower panel presents effects for boys and girls separately. This shows insignificant effects on dropout for both boys and girls.

The other outcome analyzed is being arrested. Column 2 shows that there a negative statistically significant effect of the transition on the likelihood of being arrested. Given that boys are much more likely to have been arrested than girls, it is likely that this effect is driven primarily by boys. To explore this, the lower panel presents results for boys and girls separately. As one might expect, the estimated effects are much larger for boys than for girls. For boys, the point estimate is -0.0269, which implies a 2.69 percentage point reduction in ever being arrested (p-value<0.05). The effect for girls is close to zero and is not statistically significant. Because this arrest outcome is somewhat truncated, and treated cohorts are younger than untreated cohorts, I also defined the outcome as having been arrested by the age of 16 and having been arrested by the age of 15. In columns 3 and 4, one can see that the all-boys transition led to a statistically significant reduction in the likelihood that a boy was arrested. The results for girls all have negative signs, suggesting that there could be some reduction. Overall, the results suggest a reduction in arrests for boys with some suggestive evidence of a reduction for girls.

The results thus far have focused on outcomes that occur three years after secondary school entry. However, one may wonder if there are longer-run effects that persist over time. To provide evidence on longer-run effects Table 5 presents estimated effects on the likelihood of taking advanced math, science or language courses in upper secondary school, and the likelihood of earning a secondary school completion credential. These outcomes are only available for the first post transition cohort (unless students skip a grade) such that any estimated single-sex effects are driven almost entirely by this first treated cohort. Figure 5 suggests that the estimated effect for the first cohort is likely to be similar for the other cohorts, but this is nonetheless a limitation. The first longer-run academic outcome analyzed is taking an examination for advanced level classes. Taking advanced level courses in Trinidad and Tobago is a strong indicator for likely tertiary enrollment and would also signal motivation. This signaling value is not too dissimilar to that of taking AP courses in the United States (Jackson 2010). There is some evidence that the single-sex transition increased advanced course taking. The point estimates show that boys and girls were more likely to take advanced English and math course after the transition. However, only the effect on advanced English is marginally statistically significant. Note that there is no effect on the participation margin (Column 1) so that these increases represent an increase in advanced course taking conditional on staying in school.

The last key outcome is earning a secondary school completion credential. In Trinidad and Tobago passing 5 subjects including math and English will earn one a CSEC certificate. Earning this certificate is a prerequisite to continuing in the traditional academic education system. The results show that after the transition students passed 0.41 more subjects. They were also about 4 percentage points more likely to pass the math exam and the English exam. Not surprisingly, these students were more likely to earn the certificate. The point estimate in Column 9 is 0.0493 (p-value<0.1), indicating that after the transition, students were about 5 percentage points more likely to earn a certificate—that is, 5 percentage points more likely to successfully complete secondary education. Low levels of tertiary enrollment are a problem in Trinidad and Tobago. Only 31 percent of students who enter secondary school leave with the prerequisites to enter tertiary education so that the increase is a sizable 16 percent. Overall the results indicate that the test score effects observed three years after secondary school entry are associated with increased taking of advanced courses, taking and passing more exams, and positive effects on earning a secondary school completion credential. In conjunction with the reductions in arrests, the results indicate sizable benefits to single-sex schooling for the student population under study.[28]

---

[28] The lower panel loos at the longer run effects for boys and girls separately. For most outcomes one cannot reject equality of the effects across the two groups. However, the point estimate suggest that girls are less likely to take advanced math courses in the all-girls cohorts while boys are more likely. This difference is significant at the 5 percent level. Given that all other outcomes reveal no gender differences, this is based on a single cohort, and the negative effect for girls is marginally statistically significant, I caution against over-interpreting this estimate.

IV.E. Evidence on Mechanisms

As described in Section II, single-sex effects likely reflect some combination of the three mechanisms outlined in the model (a) direct gender peer interaction effects, and the two indirect peer effects (b) the boutique effect, and (c) the focus effect. Because direct gender peer effects will always coincide with indirect peer effects, *all* peer effects papers reflect some combination of direct and indirect peer effects. Even though it is impossible to disentangle these different mechanisms, I am able to test for these mechanisms using survey data. In 2013, 2014 and 2015, I administered surveys to students in pilot schools and 20 comparisons coed schools.[29] The survey questions were designed to be sensitive to the mechanisms outlined in the model.

To identify the effect of the single-sex transition on survey responses, I rely on the assumption that students in the same school in the same grade across cohorts will be exposed to essentially the same school inputs. I *only* analyze answers to questions about classroom peers and *not about the students themselves*. Accordingly, any changes in student selection should be less of a concern. I compare survey responses of students in the same school and grade across cohorts. By looking within school grade cells, one holds school-grade level inputs constant and one also holds the level of student maturity constant. I estimate the following model by OLS.

[11] $$Q_{ijg} = All\_Boys\delta_1 + All\_Girls\delta_2 + \pi_{gj} + \varepsilon_{ijg}$$

In [11], Q_{ijg} is the response of student *i* in school *j* in grade g to the survey question, π_{gj} is a fixed effect for the school and grade so that all comparison are made within the same school and grade, and ε_{ijg} is a random error term. *All_Boys* and *All_Girls* are indicator variables equal to 1 if the school grade cell is all boys or all girls respectively in that year and zero otherwise. The parameters of interest are δ_1 and δ_2 -- the effect being in an all-boys or all-girls cohort (holding other school characteristics fixed). Standard errors are clustered at the school-grade level.

Before presenting the results, it is important to establish that the survey results are valid. The overall survey response rate was 67 percent. As such, Appendix E shows that there is no differential response by treatment status. To provide additional checks on the validity of the survey responses, I included three validation questions. The first question is whether most of the student's friends are the same gender as them, the second is whether their parents think education is important, and the third is the reported gender of the teacher. If the survey results are valid (i.e.

[29] The 20 comparison schools were chosen based on selectivity, school type and location (as were the pilot schools).

students respond as expected, there is no differential selection of students, and there were no personnel changes at treatment schools), one should see significant effects on students reporting that most of their friends are the same gender as them, we should see no effect on whether an individual's parents think education is important, and we should see no effect on teacher gender. These are the patterns observed. Taken together, these specification tests suggest that the effects on the survey responses can be taken as casual. I now present various pieces of evidence to determine which causal mechanisms are likely at play.

Are the positive effects driven by direct peer effects?

Given the positive single-sex effects for boys and girls, a simple model of linear direct peer effects where a greater proportion of boys is deleterious to the outcomes of all students cannot drive these effects in this setting. In order for direct peer effects to explain the patters of results, peer effects would have to be asymmetric such that boys have better outcomes when exposed to more boys, while girls have better outcomes when exposed to more girls (as found in Whitmore 2005, Black et al 2013; Ooserbeek and van Ewijk 2014; and Lu and Anderson 2015).

I created four indexes designed to capture the direct peer effect mechanisms presented in the gender peer effects literature. The four measures describe whether (1) peers are disruptive, (2) peers distract students from doing schoolwork, (3) students learn from their peers, and (4) peers make students anxious.[30] The estimated effects are presented in Table 6. Girls in the all-girls cohorts are less likely to report that their peers are disruptive and that their peers are distracting than girls in the coed cohorts from the same school in the same grade. These reductions are statistically significant. In contrast, boys in the all-boys cohorts are more likely to report that their peers are disruptive and that their peers are distracting, but these effects are small and not statistically significant. These patterns are consistent with Lavy and Schlosser (2012) and Hoxby (2001) that show that students in classrooms with a higher fraction of boys (on the margin) are more disruptive than those with a smaller share. However, the point estimates suggest that girls may be more sensitive to these disruptions than boys. There is also evidence that more male peers is bad in terms of direct learning from peers; girls are more likely to report learning from their peers while boys are less likely to report learning from their peer in the single-sex cohorts.

While boys learn less from their peers and report higher levels of disruption when there are more boys, there is some evidence of positive direct effects for males in other dimensions.

[30] See Appendix F for detailed discussion of the individual survey questions used to create these measures.

Specifically, boys are less likely to report that their peers make them anxious. While the results indicate unambiguously positive peer effects for girls, the overall direct peer effects for boys is ambiguous. If the benefits to being less anxious around boys is larger (or smaller) than the benefits to being in a less disruptive classroom with more girls, the direct peer effect for boys would be positive (or negative).[31] I now explore the possible role of indirect peer effects.

Are the effects driven by the boutique or the focus effect?

The other hypothesized mechanisms behind a single-sex schooling effect are indirect peer effects driven by changes in teacher instructional practices that may lead to more effective instruction or alignment of instruction to the particular needs of each sex. While I am unable to observe the extent to which instruction is aligned toward boys or girls, I am able to test for changes in students reports of teachers' instructional practices. If there is a change in alignment, there should be a change in teacher instructional practices. Even though, I do not have a complete list of instructional practices, I do have student survey data that ask students about their teachers' practices. If there are different changes in teacher practices between the all-girls and all-boys environments, it would be evidence of the boutique effect. If there are similar changes in teacher practices in the all-girls and all-boys environments, it would be evidence of the focus effect.

Because there are several survey questions, I created six indexes that measure teacher practices. The first is a measure of the extent to which teachers give students individual attention. The second index captures whether teachers use examples in their instruction. The third is the extent to which teachers attempt to monitor and track students understanding during class. The fourth is whether teacher are warm toward students, the fifth is whether teachers involve students in instruction, and the sixth is whether the teacher is strict.[32] Each measure is standardized to be mean zero with a standard deviation of 1. The regression results in Table 6 indicate that both boys and girls report higher levels of individual teacher attention in the single-sex cohorts. The fact that there is more reported teacher attention in both all-boys and all-girls cohorts is indicative of an indirect peer effect, and is suggestive of the focus effect. In fact, under the focus effect there are some efficiency gains associated with having a more homogenous student population. These efficiency gains may allow teachers to provide more individualized attention as reflected in the

[31] Note that boys do not report much higher levels of disruption in the all-boys cohorts so that the negative direct gender peer effects through increase disruption for boys may be small. That is, even if having more boys is disruptive, boys do not perceive it as such and may not be harmed by it.

[32] See Appendix F for details on the individual survey questions used to create these measures.

surveys. This is generally supported by the fact that both boys and girls report higher levels of teacher warmth in the single-sex cohorts (only the effect for boys is statistically significant). The evidence of the boutique effect (i.e. using different instructional practices in the all-girls and all-boys environments) is weak. There are no statistically significant effects on teachers using examples in class, on teachers tracking student understanding, or in whether teachers are strict. The one practice for which there is a change is the extent to which teachers involve students. Teachers appear to be marginally, more likely to involve students in the all-girls cohorts.

In sum, the results suggest that single-sex environment provide positive direct peer interaction effect for girls, while the sign of the direct per interactions for boys in ambiguous. However, the results reveal that direct peer interaction may only be a part of the overall single-sex schooling effect. Looking at indirect peer effects through teacher behaviors, the results are consistent with some efficiency gains to being in a single-sex classroom that allows teachers to give students more individualized instruction and exhibit greater warmth (i.e. the focus effect). However, (surprisingly) the improved student outcomes do not appear to be driven by greater alignment of instructional practices to the particular needs of each sex.

V. Discussion and Conclusions

The merits of single-sex education have been debated for decades in the United States and also other nations. Proponents of single-sex education view it as a panacea for a host of gender related education problems such as female underrepresentation in math and science fields and male behavior problems. Opponents argue that there is no benefit to educating boys separately from girls and highlight that there may be some deleterious socialization effects associated with being only with one's own sex. Due to the difficulty of (a) identifying the effects of single-sex schools from that of the characteristics of the individuals who attend these schools, and (b) identifying the effects of single-sex instruction from that of the characteristics of those schools that tend to be single-sex, the existing research literature has not been able to provide conclusive evidence on the matter. This paper presents a context and empirical approach that allows one to credibly addresses both of these empirical challenges. By exploiting exogenous variation to schools within a cohort (due to student assignment rules) I am able to isolate the effect of schools from that of students. By exploiting changes in single-sex status within schools over time, I am able to isolate the effect of single-sex instruction from that of any characteristics of the schools that happen to provide single-sex instruction. My results indicate that both sources of bias are important and that a failure to account for both sources of bias may lead erroneous estimates.

The results show that single-sex education can improve both boys' and girls' outcomes. Three years after being assigned to a single-sex secondary school, both boys and girls have higher scores on standardized tests. Five years later, they are more likely to take and pass advanced courses. In the long run, both boys and girls are more likely to have completed secondary school and to have earned the credential required to continue to tertiary education. Importantly, boys are also less likely to have been arrested. Taken as a whole, the results suggest that being in the single-sex cohorts improved test scores and also improved longer-run non test score outcomes such as advanced course taking, high school completion and engaging in criminal activity.

One limitation of existing studies on single-sex education is the lack of evidence on mechanisms. This paper presents some new results and analysis in this regard. First the theoretical framework highlights the different pathways through which single-sex education may affect outcomes and also makes the important point that the effect of single-sex education will depend on the nature of direct gender peer effects and also indirect peer effects that operate thought the behaviors of teachers in both single-sex and coed environments. Indeed, the survey results suggest both direct and indirect peer effect mechanisms are important.

Note that these benefits to single-sex instruction were achieved at zero additional financial cost. The test score effect of 0.16 standard deviations is about as large as the effect of going from a teacher at the 6th percentile of teacher quality to one at the 50th percentile of teacher quality. To achieve equivalent results through increases in school spending, reductions in class size, tutoring, or other interventions would require a nontrivial financial outlay. The results of this study illustrate the potential cost-effectiveness of leveraging peer effects (both direct and indirect) to improve student outcomes (both educational and otherwise). Even though the evidence presented here may not generalize to all contexts, the evidence demonstrates that single-sex education can be an effective low-cost way to improve student outcomes. Further work will need to be done to better identify the contexts in which single-sex instruction is likely to improve the outcomes of both girls and boys. The theoretical framework and the evidence on mechanisms presented will hopefully help guide future research in these areas.

Works Cited

1. Booth, Alison L., Lina Cardona-Sosa, and Patrick Nolen. 2013. "Do Single-Sex Classes Affect Exam Scores? An Experiment in a Coed University." *IZA Discussion Paper Number 7207.*
2. Black, Sandra E. & Paul J. Devereux & Kjell G. Salvanes, 2013. "Under Pressure? The Effect of Peers on Outcomes of Young Adults," Journal of Labor Economics, University of Chicago Press, vol. 31(1), pages 119 - 153.
3. Chade, H., Smith, L., 2006. Simultaneous search. Econometrica 74 (5), 1293–1307.
4. Chavous, Kevin P. 2013. "Single-Sex Education: A Viable Option for Today's Students." *The Huffington Post.* 5 4. Accessed 3 5, 2015. http://www.huffingtonpost.com/kevin-p-chavous/singlesex-education-a-via_b_3015145.html.
5. Deming, David J., Justine S. Hastings, Thomas J. Kane, Douglas O. Staiger "School Choice, School Quality and Postsecondary Attainment". 2014. American Economic Review, 104(3): 991-1014.
6. Dee, Thomas S., Brian A. Jacob, Jonah E. Rockoff, and Justin, McCrary. 2011. "Rules and Discretion in the Evaluation of Students and Schools: The Case of the New York Regents Examinations." *Columbia University Mimeo.*
7. Ding, Weili, Lehrer, Steven F., 2007. Do peers affect student achievement in China's secondary schools? The Review of Economics and Statistics 89 (2), 300–312.
8. Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* (101) 5: 1739-74.
9. Entwisle, Doris R., Karl L. Alexander, and Lina S. Olson. 2007. "Early Schooling: The Handicap of Being Poor and Male." *Sociology of Education* 114-138.
10. Fryer, Roland G. 2014. "Injecting Successful Charter School Strategies into Traditional Public Schools: Evidence From Field Experiments." *Quarterly Journal of Economics* 129 (3): 1355-1407.
11. Fryer, Roland G., and Steven D. Levitt. 2010. "An Empirical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economics* (American Economic Association) 2 (2): 210-40.
12. Gale, David, and Shapley, Lloyd, 1962. College admissions and the stability of marriage. Am. Math. Mon. 69 (1), 9–15.
13. Hardwick Day 2008 "What Matters in College After College: A Comparative Alumnae Research Study"
14. Heckman, James J., and Tim Kautz. 2012. "Hard evidence on soft skills." *Labour Economics* 19 (4): 451-464.
15. Heckman, James. 1999. "Policies to Foster Human Capital." *NBER Working Paper No 7288.*
16. Hoxby, Caroline M. 2000. "Peer effects in the classroom: Learning from gender and race variation." *NBER Working Paper no. 7867.*
17. Hoxby, Caroline and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Presented at the 2006 American Economics Association Annual Meetings.
18. Howell, William, Martin West and Paul Peterson, "The 2008 Education Next–PEPG Survey" Education Next Fall 2008, pp. 12-26.
19. Jackson, C. Kirabo. (2010) "Do Students Benefit From Attending Better Schools?: Evidence From Rule-based Student Assignments in Trinidad and Tobago" *The Economic Journal*, 120(549): 1399-1429.
20. Jackson, C. Kirabo. 2013. "Can Higher-Achieving Peers Explain the Benefits to Attending Selective Schools?: Evidence from Trinidad and Tobago." *Journal of Public Economics.*
21. Jackson, C. Kirabo. 2013. "Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina." *NBER Workign Paper No 18624.*
22. Jackson, C. Kirabo. 2012. "Single-sex schools, student achievement, and course selection: Evidence

from rule-based student assignments in Trinidad and Tobago." *Journal of Public Economics* 96 (2012): 173-187.
23. Jacobs, Jerry A. 1996. "Gender Inequality and Higher Education." *Annual Review of Sociology* 22: 153-185.
24. Ku, Hyejin, and Do Won Kwak. 2013. "Together or Separate: Disentangling the Effects of Single-Sex Schooling from the Effects of Single-Sex Schools." *University of Queensland Mimeo.*
25. Lavy, Victor, and Analia Schlosser. 2012. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics* 3 (2): 1-33.
26. Lee, Soohyung, Lesley J. Turner, Seokjin Woo, and Kyunghee Lim. 2014. "All or Nothing? The Impact of School and Classroom Gender Composition on Effort and Academic Achievement." *NBER Working Paper No 20722.*
27. Lu, Fangwen, and Michael L. Anderson. 2015. "Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups". *Journal of Labor Economics* 33 (1).: 91–122.
28. McCrary, Justin, (2008), Manipulation of the running variable in the regression discontinuity design: A density test, *Journal of Econometrics*, 142, issue 2, p. 698-714.
29. Mael F., A. Alonso, D. Gibson, K. Rogers, and M. Smith. 2005. Single-sex Versus Coeducational Schooling: A Systematic Review. Washington, DC: US Department of Education.
30. Mann, Allison, and Thomas DiPrete. 2013. "Trends in Gender Segregation in the Choice of Science and Engineering Majors." *Social Science Research* 42: 1519-1541.
31. Oosterbeek, Hessel., Reyn van Ewijk "Gender peer effects in university: Evidence from a randomized experiment" Economics of Education Review Volume 38, February 2014, Pages 51–63
32. Park, Hyunjoon, Jere R. Behrman, and Jaesung Choi. 2013. "Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools." *Demography* 50 (2): 447-469.
33. Pop-Eleches, Cristian and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses."*American Economic Review*, 103(4):1289-1324.
34. Ready, D. D., L. F. LoGerfo, D. T. Burkam, and V. E. Lee. 2005. "Explaining girls' advantage in kindergarten literacy learning: Do classroom behaviors make a difference?" *Elementary School Journal* 21-38.
35. Sharma, Amitabh. 2006. "University of the West Indies (UWI) moves to address gender imbalance - Institution initiates mentoring programmes." *Jamaica Gleaner*, December 18.
36. Steffensmeier, Darrell, and Emilie Allan. 1996. "GENDER AND CRIME: Toward a Gendered Theory of Female Offending." *Annual Review of Sociology* 459-87.
37. The White House. 2012. *Race to the Top High School Commencement Challenge.* Accessed 3 3, 2015. http://www.whitehouse.gov/issues/education/k-12/commencement-challenge.
38. The World Bank. 2014. *Ratio of female to male tertiary enrollment.* Accessed 3 3, 2015. http://data.worldbank.org/indicator/SE.ENR.TERT.FM.ZS.
39. U.S. Department of Education. 2014. *Digest of Education Statistics.* National Center for Educational Statistics.
40. Whitmore, Diane, "Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment," American Economic Review: Papers and Proceedings, 2005, 95 (2), 199-203.

Tables and Figures

Table 1: *Summary Statistics*

	(1)	(3)	(4)	(2)	(5)	(6)
	All Boys Taking the SEA	Boys at All-Boy's Pilot Schools Pre-Transition	Boys at All-Boy's Pilot Schools Post-Transition	All Girls Taking the SEA	Girls at All-Girl's Pilot Schools Pre-Transition	Girls at All-Girl's Pilot Schools Post-Transition
Assignment Variables						
Assigned to Pilot School	0.203	0.92	0.976	0.194	0.955	0.979
	(0.402)	(0.271)	(0.153)	(0.395)	(0.207)	(0.143)
Incoming Characteristics						
Times Repeated SEA	0.058	0.054	0.062	0.055	0.063	0.045
	(0.234)	(0.225)	(0.241)	(0.228)	(0.244)	(0.207)
Std. Total SEA Score	-0.164	-0.286	-0.501	0.174	-0.418	-0.262
	(1.057)	(0.826)	(0.814)	(0.923)	(0.806)	(0.737)
NCSE Outcomes						
Std. Math NCSE Score	-0.122	-0.459	-0.589	0.218	-0.201	-0.071
	(0.977)	(0.86)	(0.878)	(0.955)	(0.886)	(0.823)
Std. English NCSE Score	-0.203	-0.533	-0.644	0.295	-0.161	-0.047
	(0.967)	(0.895)	(0.951)	(0.923)	(0.948)	(0.818)
Std. Spanish NCSE Score	-0.214	-0.565	-0.607	0.287	-0.160	-0.102
	(0.957)	(0.811)	(0.914)	(0.954)	(0.919)	(0.836)
Std. Science NCSE Score	-0.122	-0.455	-0.611	0.207	-0.255	-0.147
	(0.982)	(0.831)	(0.9)	(0.956)	(0.878)	(0.795)
Std. Social Studies NCSE Score	-0.188	-0.451	-0.590	0.273	-0.141	-0.008
	(0.96)	(0.875)	(0.933)	(0.946)	(0.92)	(0.844)
Std. Arts NCSE Score	-0.144	-0.305	-0.486	0.219	0.052	-0.024
	(0.956)	(0.916)	(0.982)	(0.997)	(0.948)	(0.873)
Std. Physical Ed. NCSE Score	-0.005	-0.280	-0.369	0.084	-0.111	0.017
	(0.998)	(0.964)	(1.056)	(0.974)	(0.952)	(0.867)
Std. Technical NCSE Score	-0.152	-0.121	-0.330	0.198	0.150	0.257
	(0.928)	(0.933)	(0.942)	(1.038)	(1.075)	(0.85)
Non-Test Score Outcomes						
Drop out by age 14	0.1453	0.0709	0.1288	0.111	0.1175	0.1197
	(0.352)	(0.256)	(0.335)	(0.314)	(0.322)	(0.324)
Ever Arrested	0.036	0.048	0.013	0.005	0.013	0.004
	(0.187)	(0.214)	(0.113)	(0.073)	(0.114)	(0.062)
Arrested by Age 16	0.011	0.014	0.010	0.002	0.006	0.002
	(0.106)	(0.118)	(0.099)	(0.048)	(0.076)	(0.049)
Drop out by 16	0.2055	0.167	0.261	0.145	0.200	0.214
	(0.404)	(0.373)	(0.439)	(0.352)	(0.400)	(0.411)
Certificate	0.287	0.109	0.123	0.424	0.153	0.240
	(0.452)	(0.312)	(0.328)	(0.494)	(0.360)	(0.427)
Observations	60155	3950	5148	61445	4274	4644

Notes: This dataset is based on the population of SEA takers during years 2006 through 2012. All SEA and NCSE scores are standardized to be mean zero and unit variance in each cohort. Dropout at age 14 and age 16 are inferred based on exam participation in grade 8 and grade 10, respectively. Earning a certificate means passing five subjects in the CSEC exams (10th grade) including math and English. This is the prerequisite to entering tertiary education.
Summary statistics are provided for all boys and girls taking the SEA and also the subsample of boys and girls who attended the pilot school pre and post transition to single sex. Standard errors are presented in parentheses.

Table 2: *Effects of the All-Boys Transition in 6th Grade on Math Scores in 8th Grade*

	1	2	3	4	5	6	7	8
	Math Scores							Predicted Math Scores [a]
	OLS	OLS	OLS	OLS	2SLS	2SLS (post only)	2SLS	2SLS
Pilot (Boys) * Post	-0.287**	-0.140*	0.057	0.126*	0.158*	-0.0531	-	-
	(0.054)	(0.065)	(0.049)	(0.061)	(0.064)	(0.132)	-	-
Pilot (Girls) * Post	0.292**	0.137**	0.105**	0.130**	0.119*	0.184*	-	-
	(0.040)	(0.051)	(0.040)	(0.049)	(0.057)	(0.088)	-	-
Pilot (Girls or Boys) * Post	-	-	-	-	-	-	0.143**	0.00661
	-	-	-	-	-	-	(0.049)	(0.037)
Pilot (Boys)	-0.462**	-0.450**	-0.165*	-0.226**	-0.190+	-	-0.184+	0.0209
	(0.088)	(0.087)	(0.072)	(0.075)	(0.098)	-	(0.099)	(0.020)
Pilot (Girls)	-0.525**	-0.520**	0.0337	0.0281	0.084	-	0.074	0.0179
	(0.107)	(0.107)	(0.035)	(0.041)	(0.067)	-	(0.064)	(0.017)
Gender Fixed Effects	N	Y	Y	Y	Y	Y	Y	Y
Cohort Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Cut-off Fixed Effects	N	N	N	N	Y	Y	Y	Y
Preference Group Effects	N	N	N	Y	Y	Y	Y	Y
Cohort*SEA Score (fourth order)	N	N	Y	Y	Y	Y	Y	Y
Observations	104,228	104,228	104,228	104,226	76,345	33,145	76,345	85,809

Robust standard errors in parentheses adjusted for clustering at the assigned school level.
** p<0.01, * p<0.05, + p<0.1

Notes: Each column represents a separate regression. All models include the effect of attending a pilot school so that the coefficients on Pilot (Boys)*Post and Pilot (Girls)*Post represent the post transition change in the effect of attending an all-boys pilot school and all-girls pilot school. Column 6 only uses data after 2010 to exclude data on the pilot schools prior to the transition. The coefficient on Pilot (Girls or Boys)*Post represents the difference between the effect of attending any pilot school after transition and attending any pilot school prior to transition.

a. Predicted scores are fitted values from a regression of math scores on the number of SEA attempts (repeater status in 5th grade), the student's religion, the primary school district, selectivity of the student's primary school (a proxy for home socioeconomic status), and month of birth (to measure quarter of birth effects). Results using the predicted average of all academic subjects are similar.

Table 3: *Effects on Scores by Subject*

	Math	English	Spanish	Science	Social Studies	Arts	Physical Education	Technology Education	Academic [a]	Non-Academic [a]
	1	2	3	4	5	6	7	8	9	10
Pilot (Girls or Boys)*Post	0.143**	0.109*	0.174**	0.0950+	0.117+	0.0656	0.112	0.132	0.141**	0.107
	(0.049)	(0.055)	(0.062)	(0.055)	(0.067)	(0.113)	(0.080)	(0.110)	(0.050)	(0.083)
	11	12	13	14	15	16	17	18	19	20
Pilot (Boys)*Post	0.158*	0.0947	0.176**	0.117+	0.160*	0.143	0.108	0.125	0.158**	0.14+
	(0.064)	(0.064)	(0.065)	(0.068)	(0.071)	(0.122)	(0.100)	(0.139)	(0.059)	(0.082)
Pilot (Girls)*Post	0.119*	0.13	0.171+	0.0622	0.0513	-0.0323	0.119	0.143	0.115+	0.0545
	(0.057)	(0.090)	(0.098)	(0.071)	(0.084)	(0.147)	(0.091)	(0.127)	(0.068)	(0.124)
Pr(Girls-Boys)=0	0.489	0.96	0.716	0.446	0.334	0.325	0.941	0.981	0.428	0.512
Observations	76,345	76,360	75,781	76,246	75,016	65,052	69,825	51,575	76,360	71,754

Robust standard errors in parentheses are adjusted for clustering at the assigned school level.

** p<0.01, * p<0.05, + p<0.1

Notes: Each column in each panel (top and bottom) represents a separate regression and is indicated with a specification number (1 through 20). The dependent variable is indicated in the top row. All models include the effect of attending a pilot school so that the coefficients on Pilot (Boys)*Post and Pilot (Girls)*Post represent the post transition change in the effect of attending an all-boys pilot school and all-girls pilot school. Similarly, the coefficient on Pilot (Girls or Boys)*Post represents the difference between the effect of attending any pilot school after transition and attending any pilot school prior to transition. All models include fourth order polynomial of incoming SEA test scores for each cohort, cohort fixed effects, indicators for gender, cut-off fixed effects, and choice group indicator variables.

[a] Academic is the average of math, English, Spanish, science and social studies scores. Non-academic is the average of physical education, arts, and technology education.

Table 4: *Effects on Other Outcomes*

	Dropout NCSE	Ever Arrested	Arrested by 16	Arrested by 15
	1	2	3	4
Pilot (Girls or Boys)*Post	0.0054	-0.0169*	-0.00987*	-0.00746*
	(0.0110)	(0.0076)	(0.0047)	(0.0034)
	5	6	7	8
Pilot (Boys)*Post	0.0123	-0.0269**	-0.0123*	-0.00875*
	(0.0135)	(0.0075)	(0.0057)	(0.0040)
Pilot (Girls)*Post	-0.00478	-0.00239	-0.00634	-0.00557
	(0.0168)	(0.0082)	(0.0064)	(0.0041)
Pr(Girls effect =Boys effect)	0.44	0.01	0.42	0.52
Observations	85,816	85,816	85,816	85,816

Robust standard errors in parentheses are adjusted for clustering at the assigned school level.

** p<0.01, * p<0.05, + p<0.1

Notes: Each column in each panel (top and bottom) represents a separate regression and is indicated with a specification number (1 through 8). The dependent variable is indicated in the top row. All models include the effect of attending a pilot school so that the coefficients on Pilot (Boys)*Post and Pilot (Girls)*Post represent the post transition change in the effect of attending an all-boys pilot school and all-girls pilot school. Similarly, the coefficient on Pilot (Girls or Boys)*Post represents the difference between the effect of attending any pilot school after transition and attending any pilot school prior to transition. All models include fourth order polynomial of incoming SEA test scores for each cohort, cohort fixed effects, indicators for gender, cut-off fixed effects, and choice group indicator variables.

Table 5: *Effects on Educational Attainment Outcomes (5 Years later)*

	1	2	3	4	5	6	7	8	9
	Dropout	Adv. English	Adv. Math	Adv. Science	Sub Taken	Sub Passed	Pass Math	Pass English	Certificate
Pilot (Girls or Boys) * Post	0.000618	0.0444+	0.0156	-0.00668	0.197	0.410**	0.0408+	0.0432+	0.0493*
	(0.0214)	(0.0235)	(0.0189)	(0.0331)	(0.1580)	(0.1320)	(0.0246)	(0.0252)	(0.0200)
	10	11	12	13	14	15	16	17	18
Pilot (Boys)*Post	0.00776	0.0363	0.034	-0.00022	0.151	0.464**	0.0312	0.0485*	0.0535*
	(0.0246)	(0.0297)	(0.0235)	(0.0367)	(0.1560)	(0.1490)	(0.0241)	(0.0243)	(0.0240)
Pilot (Girls)*Post	-0.0125	0.0582*	-0.0174+	-0.0172	0.276	0.308	0.0576	0.0332	0.0411
	(0.0288)	(0.0259)	(0.0099)	(0.0422)	(0.2960)	(0.2220)	(0.0381)	(0.0498)	(0.0314)
Pr(Girls-Boys)=0	0.552	0.527	0.0397	0.71	0.688	0.519	0.503	0.77	0.739
Observations	35,874	35,874	35,874	35,874	35,874	35,874	35,874	35,874	35,874

Robust standard errors in parentheses are adjusted for clustering at the assigned school level.

** p<0.01, * p<0.05, + p<0.1

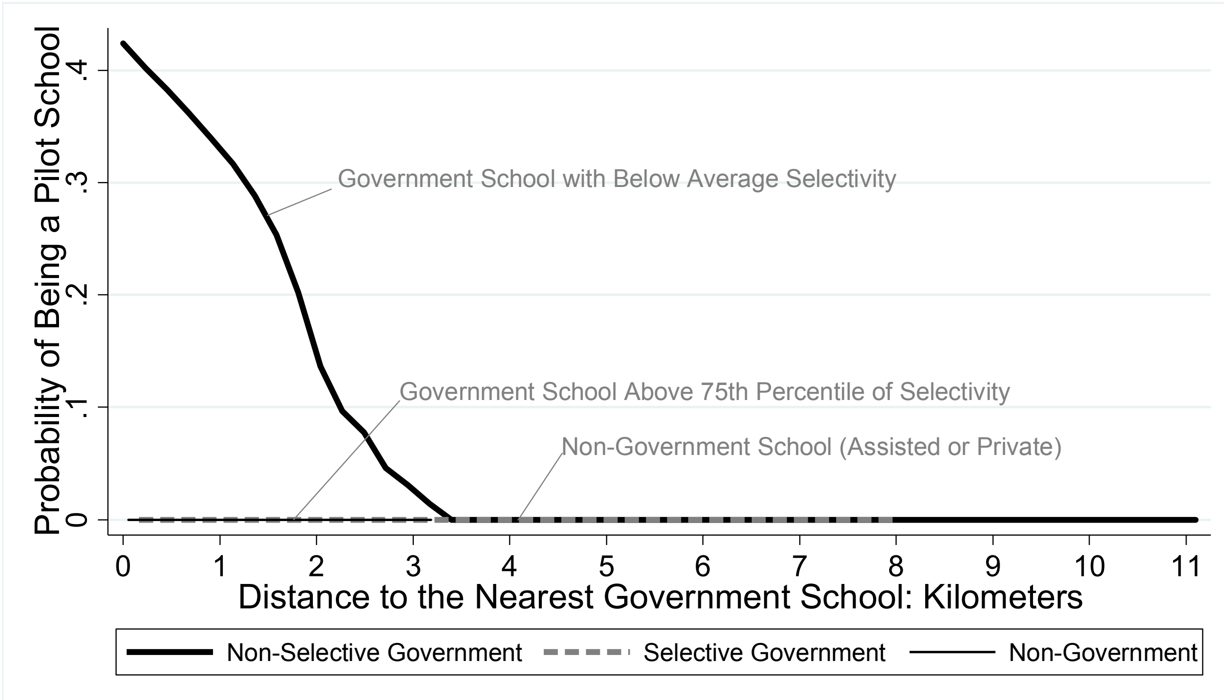
Notes: These outcomes are measured using the CSEC data from 2012-2015. As such, the regression results are only included for the SEA cohorts between 2008 and 2010. However, because some students take the CSEC exams a year early, I also include the 2011 cohort. Models without eh 2011 cohort yield almost identical point estimates, but are much less precise. Each column represents a separate regression. The dependent variable is indicated in the second row. All models include the effect of attending a pilot school so that the coefficient on Pilot (Girls or Boys)*Post represents the difference between the effect of attending any pilot school after transition and attending any pilot school prior to transition. All models include fourth order polynomial of incoming SEA test scores for each cohort, cohort fixed effects, indicators for gender, cut-off fixed effects, and choice group indicator variables.

Table 6: *Student Survey Results*

	1	2	3	4	5	6	7	8	9	10
	Peers are disruptive	Peers are distracting	Students learn from peers	Peers make students anxious	Teachers give individual attention	Teachers use examples	Teachers track student understanding	Teacher is warm toward students	Teachers involve students	Teacher is strict
Pilot (Boys)*Post	0.0357	0.0433	-0.116*	-0.107*	0.114*	0.0297	0.0115	0.128*	0.0327	0.0325
	[0.0625]	[0.0547]	[0.0504]	[0.0440]	[0.0512]	[0.0438]	[0.0432]	[0.0526]	[0.0516]	[0.0445]
Pilot (Girls)*Post	-0.194**	-0.130*	0.168**	0.0272	0.0989*	0.00381	0.0288	0.0494	0.0646+	0.0264
	[0.0593]	[0.0527]	[0.0464]	[0.0392]	[0.0462]	[0.0389]	[0.0427]	[0.0434]	[0.0386]	[0.0484]
Observations	25,250	25,250	27,948	26,596	27,845	26,538	27,239	26,378	27,554	27,991

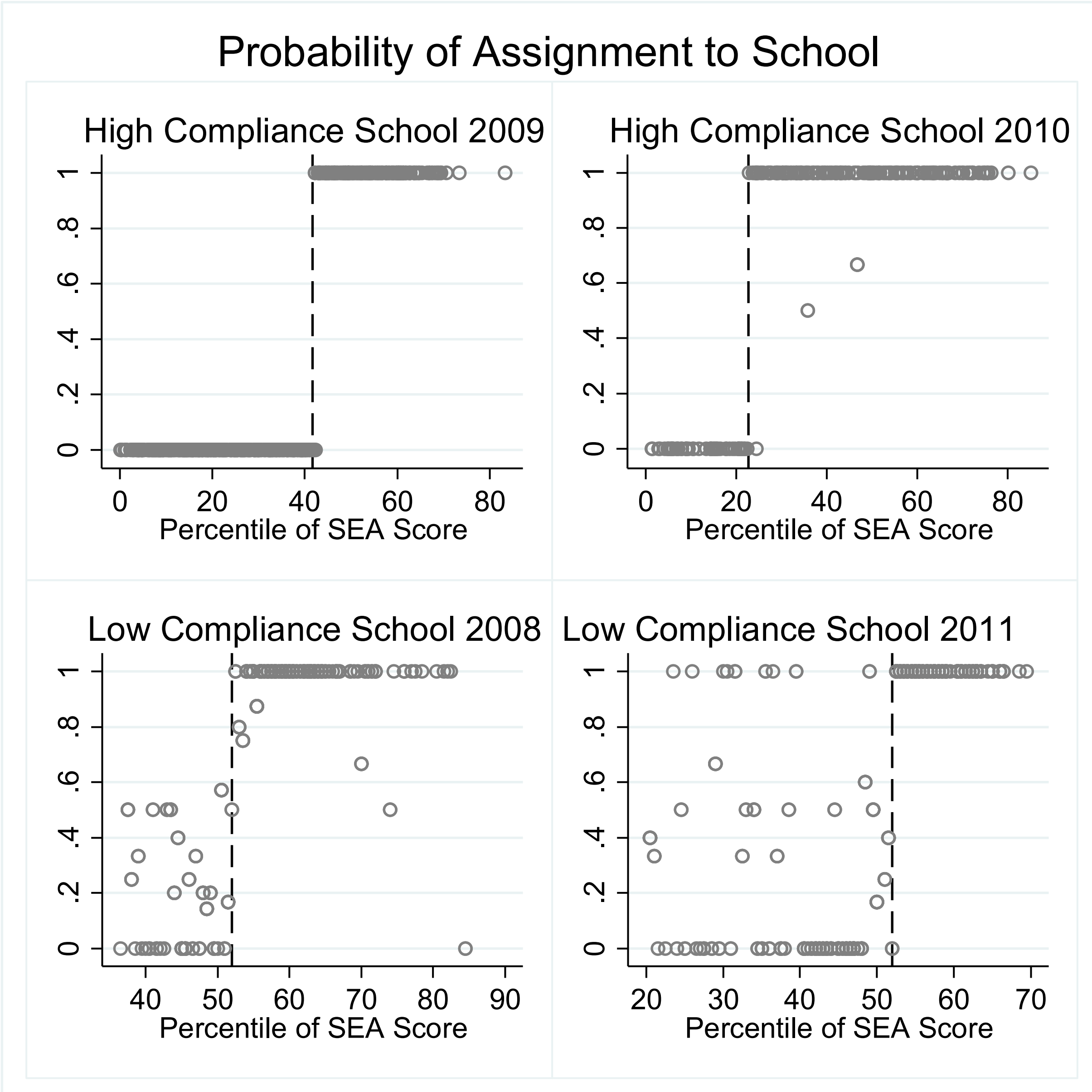
Robust standard errors in brackets are adjusted for clustering at the school-grade level.
** $p<0.01$, * $p<0.05$, + $p<0.1$
Notes: Each column represents a separate regression. The second row indicates the dependent variable. The sample is all students who attended a pilot school or one of the comparisons schools between 2012 and 2015 and also completed a survey. Because all models include school fixed effects, the coefficients on Pilot (Boys)*Post and Pilot (Girls)*Post represent the post transition change in the effect of attending an all-boys pilot school and all-girls pilot school. All models include survey year fixed effects and school-gender-form fixed effects. As such, all comparisons are made among student of the same gender at the same school (with the same teachers) in the same form but in different SEA cohorts.

Figure 1: *Predictors of Pilot School Status*



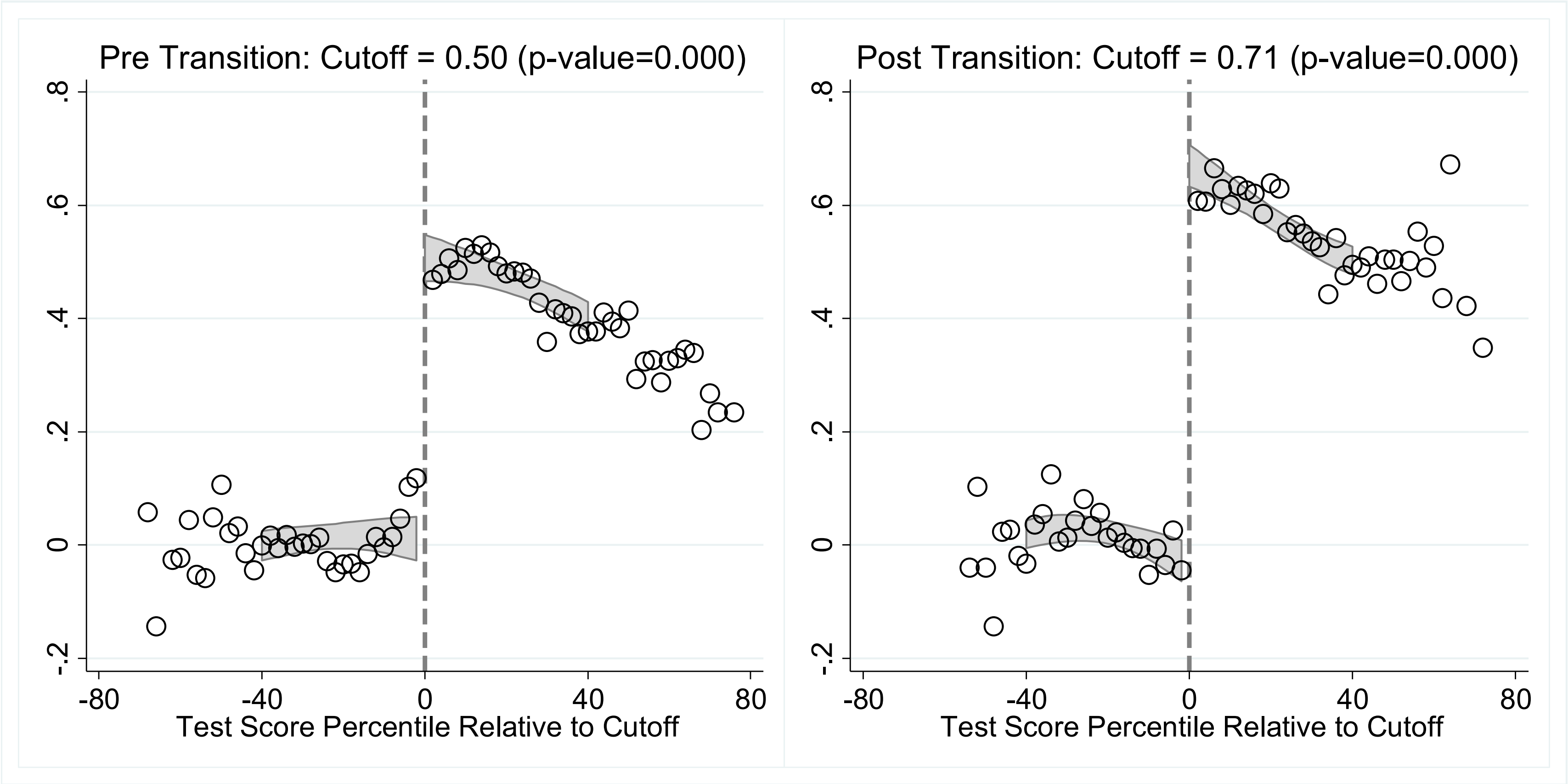
This figure depicts the probability that each school was chosen as a pilot school as a function of the distance (as the crow flies) between that school and the nearest government school. School with below average selectivity have incoming SEA scores below the mean for all schools. Schools with above 75th percentile of selectivity have incoming SEA scores above the 75th percentile for all schools.

Figure 2: *Example of cut-off based on school* <u>*assignment*</u> *for a single school in 2009 and 2010*



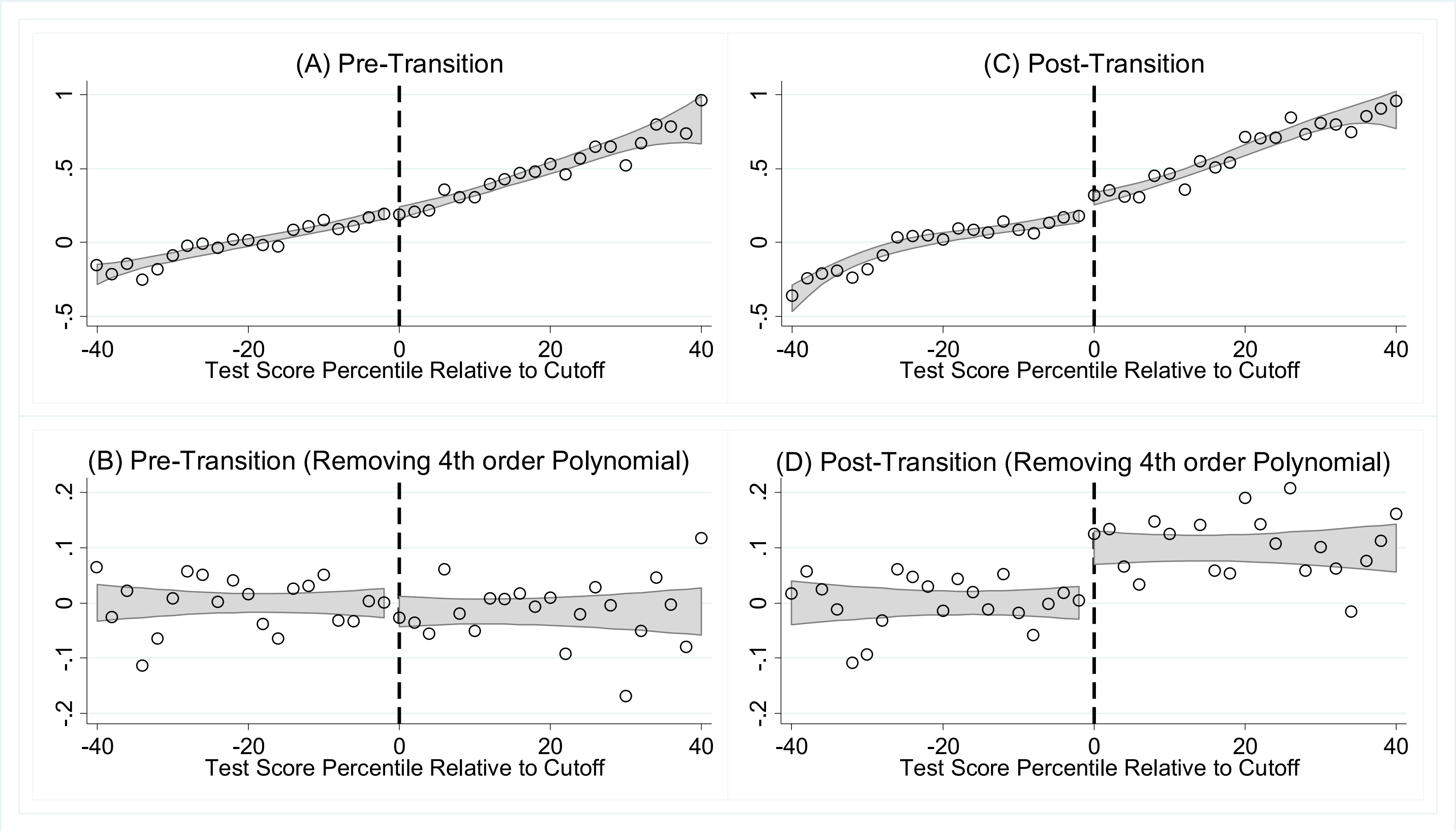
This figure depicts the probability that a student applicant to pilot school X is assigned by the MOE to that school as a function of their incoming SEA score relative to the estimated cut-off for pilot school X. the probabilities are shown for 1000 test score quantiles. The dashed line is the estimated cut-off for school X. The top two panels present data for a school with high compliance with the cutoffs (in a pre and post transition year) and the lower two panels present data for a low compliance school (in a pre and post transition year).

Figure 3: *Actual pilot attendance across pooled schools: pre and post transition*



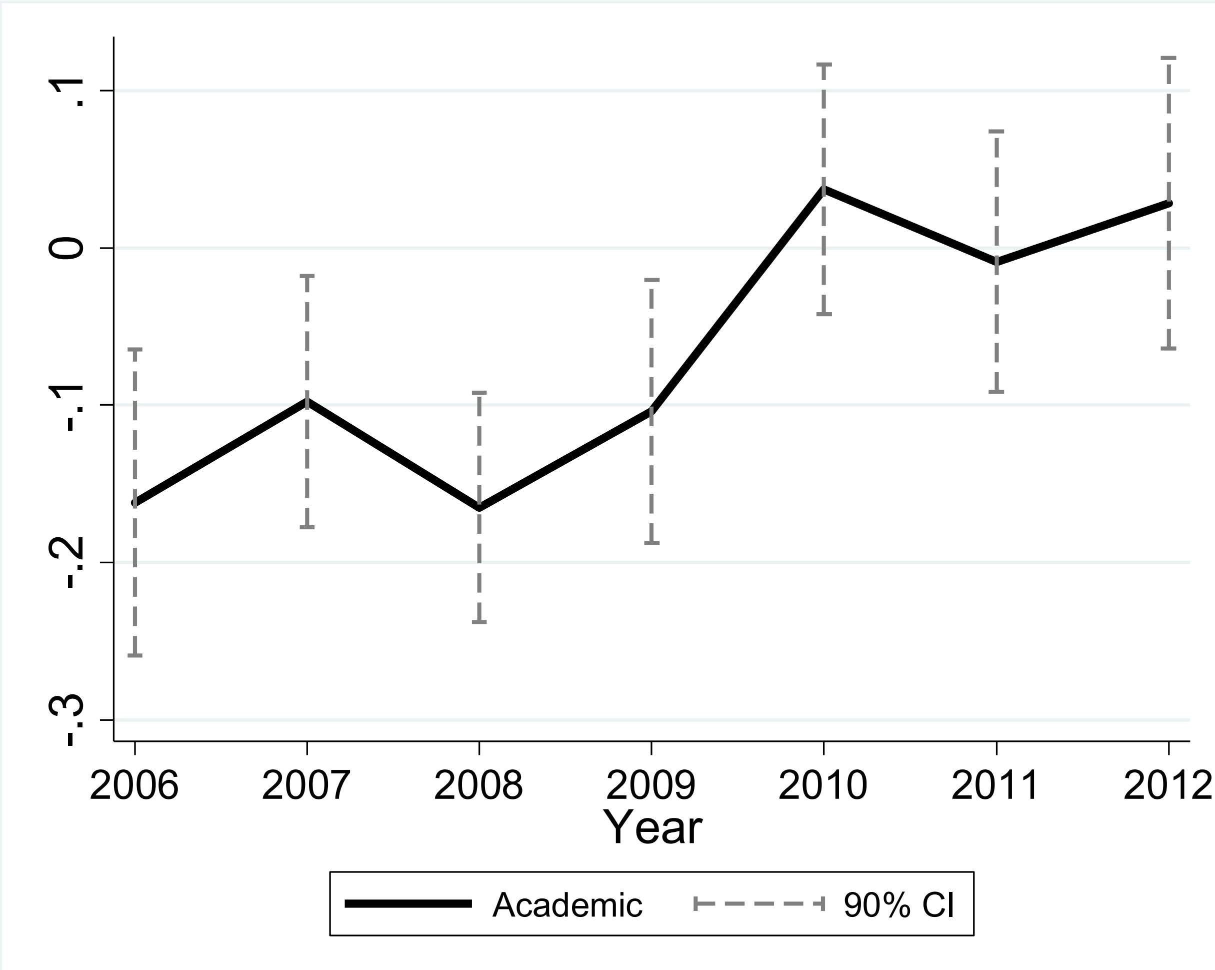
This figure depicts the probability that a student applicant to any pilot school attends a pilot school as a function of their incoming SEA score relative to the estimated cut-off for the pilot school to which they are an applicant. Following Pop-Eleches and Urquiola (2013), the data for the cutoffs for each pilot school in each year are stacked and the applicants tests scores are centered around the respective cutoff. Relative scores below 0 are below the cutoff and those above 0 are above the cutoff. The dashed grey line is the location of the estimated aggregate cut-off. The figures show the estimated outcome for each test score bin. The 95 percent confidence interval for the outcome (from a linear regression of the outcome on a fourth order polynomial in test scores and a cutoff indicator) is presented.

Figure 4: Fuzzy *RD in Math Scores across pooled schools in pre and post*



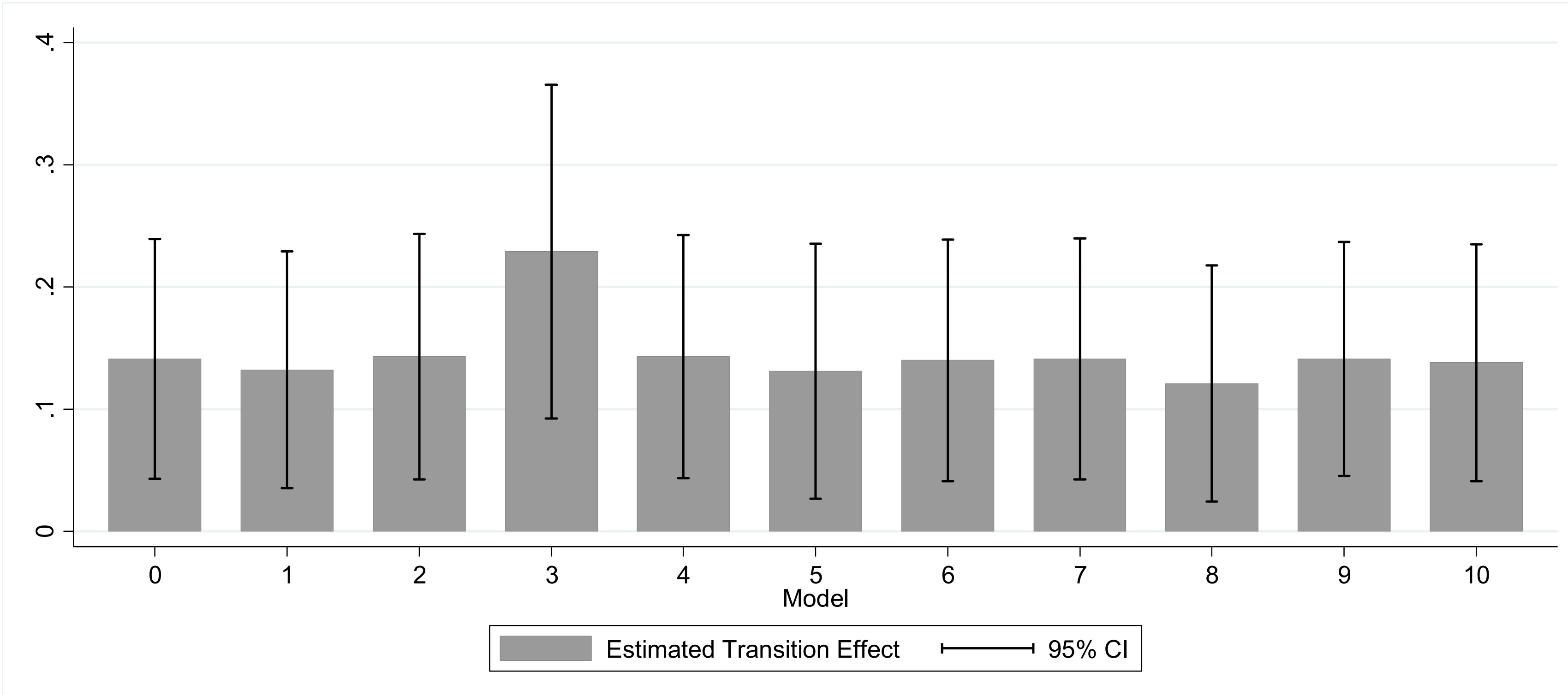
This figure depicts the average math test score as a function of their incoming SEA score relative to the estimated cut-off for the pilot school to which they are an applicant. Following Pop-Eleches and Urquiola (2013), the data for the cutoffs for each pilot school in each year are stacked and the applicants tests scores are centered around the respective cutoff. Relative scores below 0 are below the cutoff and those above 0 are above the cutoff. The dashed grey line is the location of the estimated aggregate cut-off. The top panel shows the estimated outcome for each test score bin. The 95 percent confidence interval for the outcome (from a linear regression of the outcome on a fourth order polynomial in test scores and a cutoff indicator) is presented. For a clearer presentation of the data around the cutoff, the lower panel presents the same data as the top panel bur removing smooth functions of the running variable (fourth order polynomial).

Figure 5: *RD Estimates Over Time (Event Study)*



Notes: Each data point represents the 2SLS regression estimate of attending a pilot school in each year. The endogenous variables are attending a pilot school in 2006, 2007, 2008, 2009, 2010, 2011, and 2012. The excluded instruments are indicator variables denoting scoring above the threshold for an all-male or an all-female pilot school in 2006, 2007, 2008, 2009, 2010, 2011, and 2012. All models include fourth order polynomials of incoming test score for each cohort, cohort fixed effects, gender fixed effects, and choice group fixed effects. Standard errors are adjusted for clustering at the assigned school level.

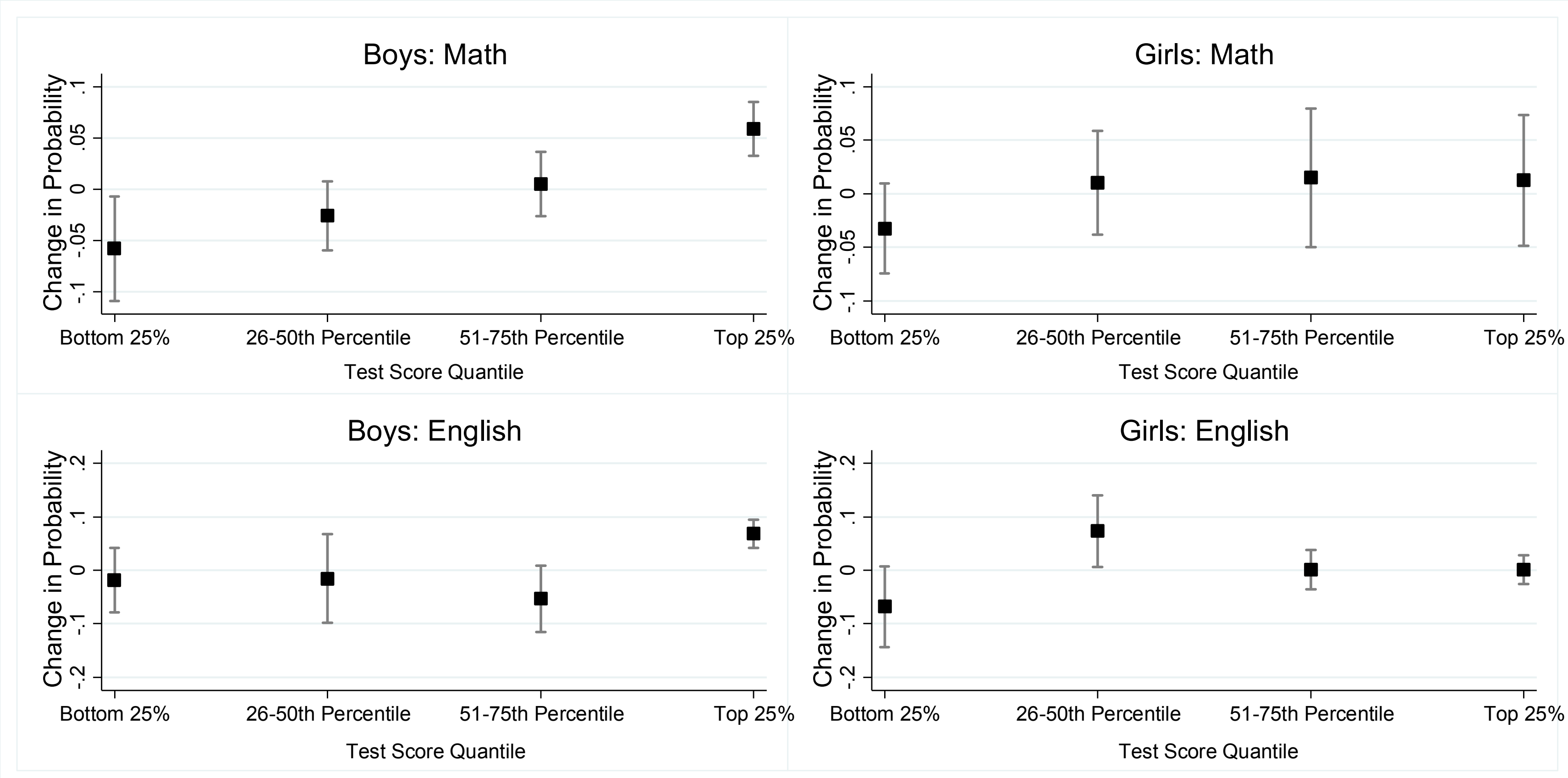
Figure 6: *Accounting for Treatment Heterogeneity in the Pilot School Effect*



Notes: Each data point represents the 2SLS regression estimate of the attending a pilot school after the transition relative to pre-transition. The endogenous variables are attending a pilot school pre and post transition. The excluded instruments are indicator variables denoting scoring above the threshold for a pilot school in 2006, 2007, 2008, 2009, 2010, 2011, and 2012. All models include fourth order polynomials of incoming test score for each cohort, cohort fixed effects, gender fixed effects, and choice group fixed effects.

Model 0 is the baseline model. Each model 1-10 includes interactions with pilot boy and girl school attendance and observed student characteristics. Model 1 includes interactions between attending a pilot school and incoming SEA scores. Models 2 and 3 include interactions between attending a pilot school and whether the students top choice school was single sex, and the number of single sex schools in the student's choice set, respectively. Models 4 and 5 includes interactions between attending a pilot school and the average SEA of the top choice school and the average selectivity of schools in the student's choice set, respectively. Model 6 includes interactions between attending a pilot school and the average SEA scores at the student's primary school. Model 7 includes interactions between attending a pilot school and the fraction of own gender peers at the student's primary school. Model 8 includes interactions between attending a pilot school and predicted math scores (predicted math scores are based on religion, primary school district, the selectivity of the student's top choice school, the average SEA scores at the students' primary school, whether they are a repeat SEA taker, and their month of birth). Model 9 includes interactions with the average selectivity of the students next ranked school (i.e. the counterfactual school), and Model 10 includes interactions with the percentile rank of the school ranked just below the pilot school.

Figure 7: *Effects on the Probability of Scoring at Various Quartiles of the Achievement Distribution: by Subject and Sex*



Notes: Within each panel, the figure presents the estimated transition effect on scoring at different quartiles of the test score distribution. Each data point represents the 2SLS regression estimate of the attending a pilot school after the transition relative to pre-transition. The 90 percent confidence interval for each estimate is presented. The endogenous variables are attending a pilot school pre and post transition. The excluded instruments are indicator variables denoting scoring above the threshold for a pilot school in 2006, 2007, 2008, 2009, 2010, 2011, and 2012. All models include fourth order polynomials of incoming test score for each cohort, cohort fixed effects, gender fixed effects, and choice group fixed effects. Each panel presents estimated effects for a difference gender and subject combination.

Appendix A: *Proofs for Claim in Theoretical Section*

In mixed-gender classrooms, some teachers may align instruction to girls all the time, others may align instruction to boys all the time, and some teachers may align instruction to boys and girls some of the time. The choice of how to align instruction will depend on the specific parameters of the production function and the incentives faced by teachers and cannot be determined *ex-ante*. I outline a three illustrative cases. In all cases, I make the assumption that teacher utility is a weighted average of the utility for each student in her class, so that $W(Y_{i\in j}) = \sum_{i\in j} l(y_i)$.

Case 1a: If the costs to adopting a mixed strategy is sufficiently large (i.e. ψ is very large), and teachers payoffs are convex in individual student test scores, then teachers will chose to align instruction only to girls. Intuitively, with very large costs to adopting a mixed strategy, teachers will align instruction to only girls or only boys even in mixed gender classrooms. If teacher payoffs are convex in individual student scores, then the marginal increase in utility will be higher for increasing tests scores at the top of the distribution than at the bottom. Because girls are more highly represented at the top of the distribution, teachers will choose only girl aligned instruction in coed classrooms.

Case 1b: If the costs to adopting a mixed strategy is sufficiently large (i.e. ψ is very large), and teacher payoffs are concave in individual student test scores, then teachers will chose to align instruction only to boys. Intuitively, with very large costs to adopting a mixed strategy, teachers will align instruction to only girls or only boys even in mixed gender classrooms. If teacher payoffs are concave in individual student scores, then the marginal increase in utility will be higher for increasing tests scores at the bottom of the distribution. Because boys are more highly represented at the bottom of the distribution, teachers will choose only boy aligned instruction in coed classrooms.

<u>Proof:</u> Assume that the distribution of female incoming achievement has the same shape as that of males, but is a right shift of that for males by some positive constant a -- this is consistent with my data. If the teacher aligns instruction to males/females, then the male/female latent outcome distribution is shifted to the right by some constant δ (i.e. $h_g p_j$=$-(h_b p_j) = \delta$). For each percentile, p, of the male distribution with latent outcome $y_{p,male}^{latent}$, there is a female with latent outcome $y_{p,male}^{latent} + a$. If the teacher aligns instruction to males then the increase in payoff for the teacher is $W\left(y_{p,male}^{latent} + \delta,\ \cdot\right) - W(y_{p,male}^{latent},\ \cdot)$ for that student. If the teacher aligns instruction to females then the increase in payoff for the teacher is $W\left(y_{p,male}^{latent} + a +\ \delta,\ \cdot\right) - W(y_{p,male}^{latent} + a,\ \cdot)$ for the analogous female student. If teacher payoffs are convex, then $W\left(y_{p,male}^{latent} + a +\ \delta,\ \cdot\right) - W\left(y_{p,male}^{latent} + a,\ \cdot\right) > W\left(y_{p,male}^{latent} + \delta,\ \cdot\right) - W(y_{p,male}^{latent},\ \cdot)$ for all percentiles of the male distribution so that the teacher's payoff is higher if she aligns instruction to females. Conversely, if teacher payoffs are concave, then this inequality is reversed for all percentiles of the male distribution so that the teacher's payoff is higher if she aligns instruction to females.

Case 3: If ψ is small, teachers may adopt some mixed approach. In the extreme case where there is no cost to adopting a mixed strategy, teacher payoffs are linear in the average for the classroom, there are equal number of boys and girls in the classroom, and boys and girls are equally responsive to alignment, teachers will be indifferent between adopting a mixed strategy or aligning instruction

to only one sex. In such cases, a mixed strategy may be adopted in coeducation classrooms.

Proof: With linear payoffs, equal numbers of boys and girls, and equal responsiveness to alignment for boys and girls, the teachers expected payoff function under female alignment minus her payoff under male alignment is $\tau[h_g(p) - h_b(p)]$, where τ is some scalar. If the average marginal effect of alignment is the same for both male and female students, then $h_g(p) = h_b(p)\ \forall\ p$, so that $\tau\big[h_g(1) - h_b(1)\big] = 0$. In such a scenario, teachers are indifferent between aligning instruction to boys only, girls only, or adopting some mixed strategy. It is easy to come up with other scenarios in which teacher will chose some mixed strategy.

Appendix B: Single-sex Status by Year and Grade in Pilot Schools

Appendix Table B1: *Status of Grades by Academic Year and Grade*

	Calendar Year Admitted Form 1 (6th grade) class						
	2009	2010	2011	2012	2013	2014	2015
Grade 6	Coed	**Single-Sex**	**Single-Sex**	**Single-Sex**	**Single-Sex**	Coed	Coed
Grade 7	Coed	Coed	**Single-Sex**	**Single-Sex**	**Single-Sex**	**Single-Sex**	Coed
Grade 8	Coed	Coed	Coed	**Single-Sex**	**Single-Sex**	**Single-Sex**	**Single-Sex**
Grade 9	Coed	Coed	Coed	Coed	**Single-Sex**	**Single-Sex**	**Single-Sex**
Grade 10	Coed	Coed	Coed	Coed	Coed	**Single-Sex**	**Single-Sex**

Appendix C: *The School Assignment Algorithm*

School slots are assigned in rounds such that the most highly subscribed/ranked school fills its spots in the first round, then the next highly subscribed school fills its slots in the second round, and so on until all school slots are filled. This is done as follows: (1) the number of school slots at each school n_j is predetermined based on capacity constraints. (2) Each student is tentatively placed in the applicant pool for her first choice school and is ranked by SEA score. (3) The school at which the n_j^{th} ranked student has the highest SEA score is determined to be the most highly subscribed/ranked school and the top n_{j1} students in the applicant pool for top ranked school j_1 are admitted to school j_1. The SEA score of the n_{j1}-th student is the cut-off score for school j_1. (4) The top ranked school slots and the admitted students are removed from the process, and the second choice becomes the new "first choice" for students who had the top ranked school as their first choice but did not gain admission. (5) This process is repeated in round two to assign students to the second highest ranked school j_2 and determine the cut-off score for the second ranked school, and this is repeated in subsequent rounds until all slots are filled. This assignment mechanism is a deferred acceptance algorithm (Gale and Shapley, 1962) in which students have incentives to truthfully reveal their rankings among chosen schools.

While the optimal set of school choices is difficult to solve, Chade, and Smith (2006) demonstrate that the choice set should include the school with the largest expected payoff (utility conditional on attendance times the likelihood of admission), students should rank selected schools in order of actual preferences, and should include a "reach" school for which admission is unlikely but the utility conditional on attendance is high.

This process is used to assign over 90% of all students. As such, as a practical matter, one can consider this applying to all students. However, there are a few exceptions to this rule. First, Government assisted schools (not analyzed in this study) are allowed to admit 20% of their incoming class at the principal's discretion. None of the pilot school is Government assisted so that there is no problem of principals hand picking students at the pilot schools. However, there are also assignments that do not follow this rule because student who do not score high enough to be assigned to a school in their choice list receive an administrative assignment from the Ministry of Education (these assignments are made to balance space considerations). Finally, due to unforeseen circumstances some schools may have less capacity than expected or may close (this may happen due to flooding etc.). In such rare cases, the Ministry will assign students to schools based on open slots in nearby schools, open slots in other school in the choice list, and proximity.

I aim to use this assignment rule to isolate exogenous variation to the 20 pilot schools. A key feature of this assignment rule is that each school has a test score cut-off above which applicants are very likely to be assigned and below which applicants are very unlikely to be assigned. Even though the cut-offs are not known to the public and not all the administrative assignments follow the cut-off rule (due to a few exceptions *made by the MOE*), because I have access to the administrative assignment (which is outside the control of the students or their parents), the cut-offs can be recovered empirically for the 20 pilot schools.

Appendix D: *Robustness Checks and Test of Validity*

The exogenous variation used in this paper is driven by the test score-cut offs. Even though there is no way to prove for certain that the cut-off variation is valid, here I present evidence that this identification strategy is valid. One key diagnostic is to test for excess density above the cut-off and less than expected density below the cu-off (McCrary 2008). The first prima facie evidence of no change in density through the cut-off is simply the histogram of relative scores. As one can see in Figure 3, there is no uncharacteristic spike in density above the cut-off or dip in density just below the cut-off. If one computes the density of observations at each relative score are regresses this on scoring above the cut-off along with smooth functions of the relative score, there is no statistically significant relationship between scoring above the cut-off and the density. The point estimate is -0.00119 (p-value=0.363) – negative and not statistically significant. Taken together, the patterns suggest that the variation due to the test score cut-offs is likely valid. The other common test is for smoothness of latent outcomes (as proxied by covariates) through the cutoffs. In presenting the results I present effects on both actual outcomes and predicted outcomes (based on covariates) that show that the cut-off variation in pilot school attendance is not associated with any changes in predicted outcomes – consistent with the cut-off variation being valid.

Section III.D shows that there is little evidence of gaming around the cut-offs regarding the density of observation at each test score. However, the validity of the design also requires that there be no sorting of students around the cut-off (i.e. that latent outcomes are smooth through the cut-off). Given that students are unaware of the location of the cut-offs and are forced to make school choices before they take the SEA examinations, it is very unlikely that there is any sorting around the test score cut-offs. However, to provide further evidence that the variation employed (due to the cut-offs) is valid, I create a predicted math score variable and test for whether the 2SLS model (using the cut-of variation) predicts any change in predicted math scores.

Specifically, I regress the NCES math score on indicator variable for religion, primary school district, the selectivity of the student's top choice school, the average SEA scores at the students' primary school, whether they are a repeat SEA taker, and their month of birth. These variables are very strong predictors of math scores such that they yield an R-squared of 0.27. I then take the fitted value from this regression as my predicted math score. If there is some gaming of the cut-off, one would likely see that attending a pilot school as a result of scoring above the cut-off should be associated with better "predicted" scores. However, with no gaming there should be no relationship between scoring above the cut-off and ones predicted score. Consistent with no gaming, there is no relationship between scoring above the cut-off and ones predicted math score. Column 7 of Table 3 shows the estimated effect on predicted math scores for the all-boys pilot schools. The coefficients on the variables of interest are both small and statistically insignificant – indicating no gaming of the cut-off to the all-boys pilot schools. Column 7 of Table 4 shows the estimated effect on predicted math scores for the all-girls pilot schools. As with the boy's schools, the coefficients on the variables of interest are both small and statistically insignificant – indicating no gaming of the cut-off to the all-girls pilot schools. Results for predicted average academic scores are very similar.

Another concern one may have is that the 2SLS effects driven by the cut-offs could have happened by random chance by choosing any arbitrary cut-offs. To assuage this concern, I assigned random "placebo" cut-offs to each pilot school and then estimated the effect on math scores of scoring above the cut-off for a pilot school after the transition (after 2010). I did this with a new random draw of placebo cut-offs 1000 times and compared the placebo estimates to the actual estimate using the actual cut-offs. The real cut-off effect is 0.1681 (p-value<0.01). Only 9 out of the 1000 placebo cut-offs was as large as the actual cut-off effect. This indicates that the estimated cut-off effect (and the associated 2SLS estimates) is much larger than on might expect by random chance if there were indeed no single-sex effect on math scores.

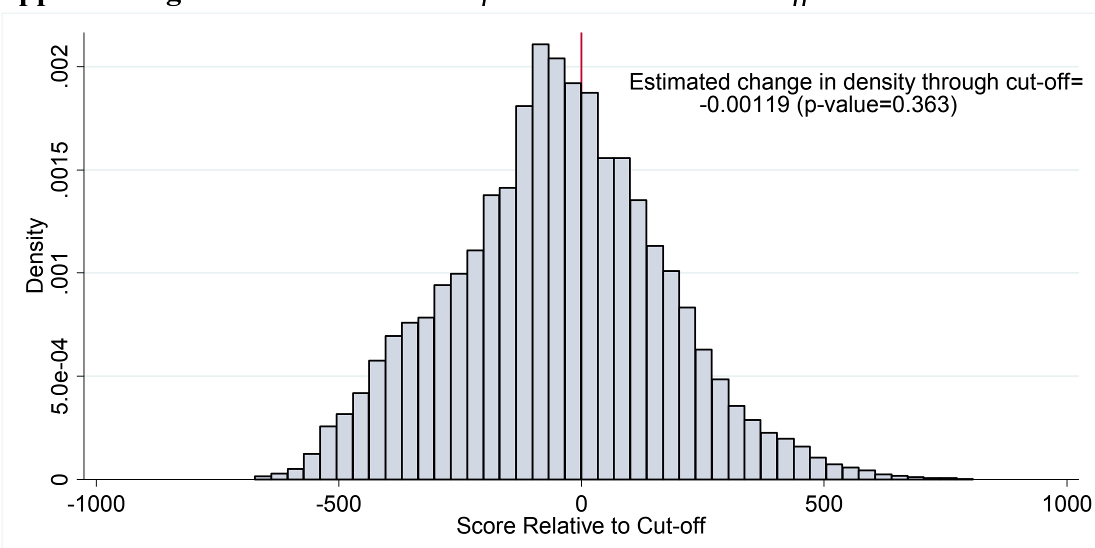
Validity of DiD

Finally, even if the RD variation removes selection to schools within a cohort, in order for the estimates to represent the causal effect of single-sex education requires that there were no changes in the pilot schools over time that coincided with the pilot program. As discussed previously, confounding changes within schools is very unlikely because (a) the individual schools had no control over when they would become pilot schools, (b) the government stipulated that no other changes take place in these schools, and (c) schools were made award of the changes during the summer preceding the change so that schools has no time to react to the change in policy before the start of the school year. Another potential problem would be that the MOE selected schools that were already on an upward trajectory to be pilot schools. To show evidence that this does not drive the results, Figure 7 presents the selection free RD estimates for attending a pilot school for 2007 through 2011. Each data point is a separate RD point estimate and the 95 percent confidence interval is also presented for each RD estimate. As one can see the RD estimates for 2007, 2008, and 2009 are all very similar to each other and hover just around -0.09. This shows that there is no indication of any upward trend for the pilot schools prior to the transition. Consistent with the positive single-sex effects for both boys and girls in Table X, the pilot school effects in 2010 and 2011 are much higher than that in the pre-treatment years. These post treatment RD estimates hover around 0.1, suggesting that the transition increased the RD pilot school effect by roughly

0.09+0.1=0.19σ. This is the same order of magnitude of the 2SLS estimates presented in Tables X and X. Importantly, Figure 7 presents a clear visual representation of the DiD variation employed and provides compelling evidence that the DiD identifying common trends assumption is valid.

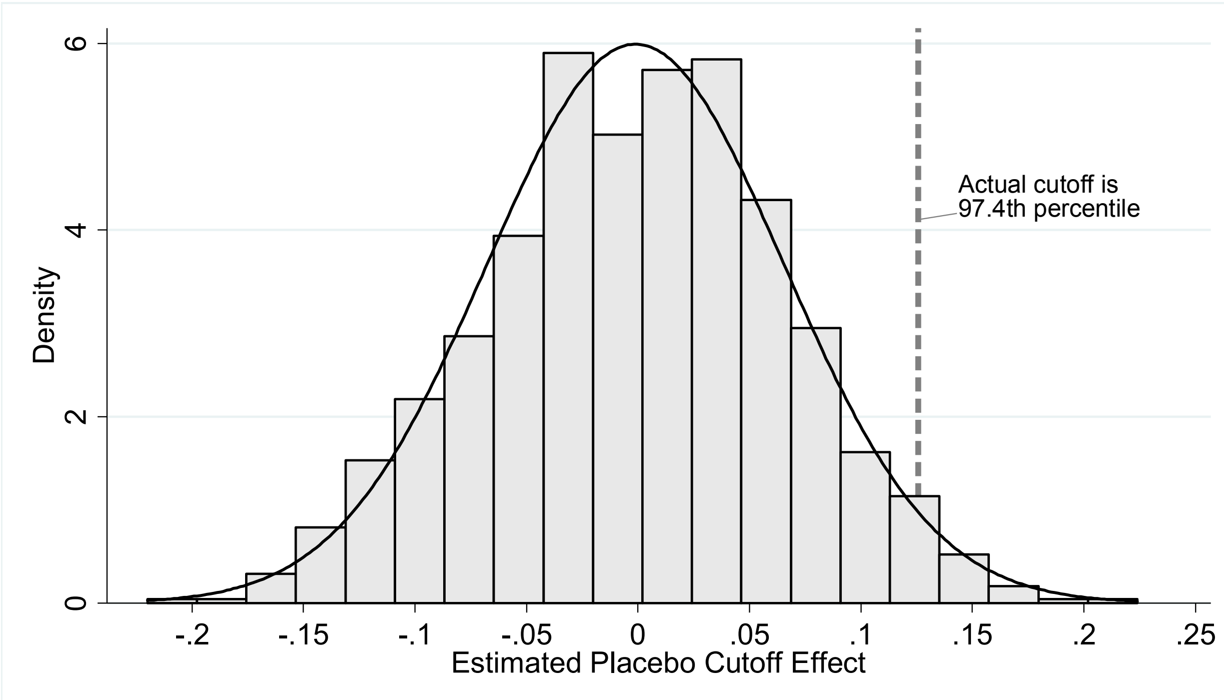
Even though the MOE clearly stipulated that there be no changes, as a final check on the DiD assumptions of no other changes, I used survey data collected by the MOE on a sample of teachers at all schools obtained in 2009 and 2013 to see if there were any systematic personal changes that occurs during the transition. In a simple difference in difference regression model predicting school level teacher characteristics with school fixed effects and year fixed effects (Appendix Table D1), between 2009 and 2013 the transition schools saw no differential change in the percentage of female teachers, or the percentage of teachers with a Bachelor's degree. There is a marginally statistically significant effect on year of birth at the all-male pilot schools. This is consistent with sampling variability and is in the opposite direction of what would be required to generate a positive transition effect for the all-male pilot schools. I also explore if differences in class size (proxied by cohort size) can explain the results. In a simple model predicting *initial* cohort size (i.e. the size of the assigned cohort) with school fixed effects and year fixed effects, the post transition pilots have no statistically significant difference in cohort size (and the effect for the all-male and all-female pilot schools are in opposite directions). The results corroborate the notion that there were no other changes at pilot schools.

Appendix Figure D1: *Distribution of Score around the cut-off*



For each of 1000 quantiles of SEA scores in each cohort, the density of observations is computed. This figure represents the density of incoming SEA test scores relative to the test score of the applicant student. The cutoff for the school to which the student has applied is 0.

Appendix Figure D2: *Transition Effect with Placebo Cutoffs*



Notes: This figure depicts the distribution of the post-pre adoption cut-off effect for the pilot schools where the cutoffs are generated randomly. In each round a random cutoff is generated for each school in each year and the estimated difference between scoring above the cut-off after the transition relative to that before is estimated. The figure depicts the distribution of the random "placebo" transition effects for the 2000 rounds. Only 32 out of the 2000 round yielded estimates as large as the real estimate.

Appendix Table D1: *Evidence of no other Changes*

	1	2	3	4
	Teacher Variables at School-Year Level			
	Female	BA Degree	Year of Birth	Cohort Size
Post*(Pilot Girl's School)	0.0243	0.108	-0.469	5.534
	[0.0473]	[0.131]	[1.325]	[10.75]
Post*(Pilot Boy's School)	0.0287	-0.0254	1.413+	-12.36
	[0.0279]	[0.177]	[0.788]	[14.53]
Years		2008 & 2013		2006-2012
Observations	240	240	240	1,060

Robust standard errors in brackets

** p<0.01, * p<0.05, + p<0.1

These models are simple difference in difference regression models that include school fixed effects, year fixed effects, and an indicator for being a pilot school post transition (Post*(Pilot Girl's School) and Post*(Pilot Girl's School)). Models 1,2 and 3 are based on a survey on teachers administered in 2008 and 2013. Column 4 is based on the administrative SAE data for years 2006 through 2012. Cohort Size is the admitted cohorts size.

Appendix E: *Student Survey Specification Checks*

Appendix Table E1: *Student Survey Specification Checks*

	1	2	3	4
	Survey Participation Rate (Cohort Level)	My Parents think Education is Important	Most of my friends are the same gender as me	Teacher is Female
Pilot (Boys)*Post	-0.0317	0.00476	0.299**	-0.0202
	[0.0356]	[0.0496]	[0.0527]	[0.018]
Pilot (Girls)*Post	-0.00231	0.0117	0.243**	0.0058
	[0.0423]	[0.0363]	[0.0587]	[0.0175]
Survey Year Effects	Y	Y	Y	Y
School-Gender-Form Effects	N/A	Y	Y	Y
Observations	609	27,477	27,514	25,886

Robust standard errors in brackets

** p<0.01, * p<0.05, + p<0.1

Each column represents a separate regression. The sample is all students who attended a pilot school or one of the comparisons schools between 2012 and 2015 and also completed a survey. Because all models include school fixed effects, the coefficients on Pilot (Boys)*Post and Pilot (Girls)*Post represent the post transition change in the effect of attending an all-boys pilot school and all-girls pilot school. All models include survey year fixed effects and school-gender-form fixed effects. As such, all comparisons are made among student of the same gender at the same school (with the same teachers) in the same form but in different SEA cohorts.

Appendix F: *Student Questions used to Construct Indexes*

Appendix Table F1: *Student Questions About Peers*

	Peers			
Survey Question	Disruptive peers	Distracting peers	Learn from Peers	Nervous around peers
My classroom is orderly	0.4749	-0.0619		
I feel safe in the classroom	0.4083	-0.0541		
My classmates distract me from my schoolwork	-0.0697	0.3345		
We do not waste time in my classes	0.3506	-0.1319		
My classmates sometimes encourage me to misbehave	-0.1112	0.3252		
I learn from my classmates			1	
I worry about impressing classmates while in class				0.4548
I am afraid to raise my hand in class				0.4748
I get nervous when I am compared with classmates				0.3498

Notes: This table presents the factor loadings for each survey question used to construct each factor. The factor is listed at the top of each column. The individual survey items are listed in the rows.

Appendix Table F2: *Student Questions about Teachers*

	Teacher Related Questions					
	Spend one-on-one time	Use examples to relate topics	Check for understanding	Warmth toward students	Involves students	Teacher is strict
Survey Question						
My teachers pay attention to me in class	0.4839					
My teachers spend one-on-one time with me in class	0.4839					
Teachers use many examples that help us understand		0.5344				
Teachers ask us for several ways to solve the same problem		0.6048				
The teachers require us to relate different topics		0.521				
Teachers check whether we know the material well			0.6148			
Teachers give assignments that help us memorize material			0.621			
Teachers hold discussions that help us understand things			0.5791			
Teachers ask us to explain our answers			0.4202			
My teachers like me				0.5572		
My teachers care about me				0.5777		
My teachers often make me feel bad				-0.2335		
Teachers praise my efforts				0.5149		
My teachers listen to my ideas					0.4996	
My teachers involve students in making decisions					0.4996	
My teachers are strict						1

Notes: This table presents the factor loadings for each survey question used to construct each factor. The factor is listed at the top of each column. The individual survey items are listed in the rows.