The Experiences of One New York City High School Cohort: Opportunities, Successes, and Challenges

Douglas Ready, Thomas Hatch, Miya Warner & Elizabeth Chu
Teachers College, Columbia University

October 8th, 2013

THE EXPERIENCES OF ONE NEW YORK CITY HIGH SCHOOL COHORT: OPPORTUNITIES, SUCCESSES, AND CHALLENGES

Douglas Ready, Thomas Hatch, Miya Warner, & Elizabeth Chu
Teachers College, Columbia University

Introduction

This paper examines the experiences of a single cohort of New York City public high school students who entered ninth grade in 2005 and who hoped to graduate in 2009. During this period New York City schools showed strong improvements, and the outcomes associated with these 77,501 students were clearly representative of these positive trends.[1] For example, graduation rates grew from 51% for students entering ninth grade in 1999 to 67.1% for the student cohort we examine here. However, many students in this cohort continued to encounter individual challenges and structural barriers that impeded their social and academic development, and our work points to these as well. Indeed, many findings throw into stark relief historic inequities that remain major challenges today. Our aim here is to unify a relatively distinct (though clearly related) set of analyses into a single document to foster dialogue and advance discussions about the public schools in New York City[2]

Structure of the Paper

Although the focus of the three papers in this commissioned series is career and college readiness, it is worth noting that it was only as the members of this cohort were entering middle school that "college and career ready" started to become a focus for educational discussions and policymaking. While the demand to prepare all students for college and careers may not seem "new," Leslie Siskin makes clear in the first paper in this commissioned series that, in fact, this *is*

[1] Of the 77,501 students who began ninth grade in 2005 and were projected to graduate in Spring 2009, over 40% were Hispanic and roughly 35% were black, 13.2% were white and 10.8% Asian, and a small number of students were identified as Native American or multi-racial. Roughly half received free or reduced-price lunches, and almost one in six were identified as having special needs. The cohort was also linguistically quite diverse, with 12% classified as Limited English Proficient. Spoken languages ranged from Afrikaans and Amharic to Yiddish and Yoruba.

[2] Readers should bear in mind that this student cohort began kindergarten roughly seven years prior to the Bloomberg era. As such, much of their important early schooling occurred in a reform context quite different from that in place today.

a new demand. Illustrating this growing demand, in 1980, only about 40% of high school sophomores aspired to post-secondary education. By the time this New York City student cohort entered high school in 2005, the National Survey of High School Student Engagement reported that 90% of high school students expected to go to college (Siskin, 2013).

Since this focus on college and career is recent, there are few "college and career ready" performance measures available for this cohort. However, our data do permit us to investigate college and career preparation across several important dimensions. In the first section of this paper we provide a brief overview of this cohort's K-8 experiences. Discussions about career and college readiness typically focus on high schools. Often lost, however, is the fact that students enter high school with dramatically different academic backgrounds. Whether students leave high school with the requisite skills is partly a function of their experiences in elementary and middle school. Students who begin high school with weak academic backgrounds are far less likely to graduate, much less be career and college ready. Despite this variability in students' prior educational experiences, New York City high schools are now expected to graduate every student and ensure that their diverse student populations are college and career ready—a challenging task, indeed. Given the importance of students' K-8 experiences, this paper begins with analyses that touch on student characteristics and school contexts that influence later success in high school. Specifically, we explore student academic performance, attendance, and behavior, and describe the extent to which these outcomes varied across student sub-groups. We also examine school demographic and academic compositions and how they changed as students progressed through the system, and the school-to-school pathways these students took toward high school.[3]

In the second section we describe a second set of analyses that focus on high school outcomes and post-secondary enrollment and persistence through the first year of college. Given that most would consider high school completion to be a minimal prerequisite for college and career readiness, the first set of high school analyses addresses high school graduation. We then focus more explicitly on post-secondary preparedness by exploring the extent to which students in this cohort encountered rigorous coursework. We concentrate on mathematics coursetaking

[3] These analyses explore achievement gaps within the single student cohort over time. They do not explore how these gaps changed across cohorts over time.

and students' progression through the mathematics curriculum—a crucial predictor of later college success.[4]

The third and final section of the paper focuses directly on post-secondary outcomes, and explores student social and academic background characteristics that predicted college enrollment and persistence through the end of the first year of college. Technical details are presented in footnotes or appendices. Readers interested in more specific details are encouraged to contact the first author.

Part 1: K-8 Outcomes

Academic Outcomes

We began by exploring whether students who had struggled academically in elementary school had caught up by the start of high school. Despite some positive movement, only 2.7% of students who failed to meet the third-grade English Language Arts (ELA) standard went on to meet or exceed the ELA benchmark in eighth grade, and only one in three of the students who failed to meet the third-grade ELA standard graduated from high school. Conversely, 91.3% of students who exceeded the ELA standard in third grade would meet or exceed the standard in eighth grade, and almost 90% of these students graduated within four years. We found similar results for mathematics. In short, early academic performance was strongly tied to later success, and to a remarkable degree, students in this cohort maintained their academic positional advantages or disadvantages from elementary and middle through the start of high school.

This cohort also began high school with substantial racial/ethnic disparities in academic performance that can be traced back to at least third grade.[5] The average white third grader scored at the 75th percentile on the ELA, while black and Hispanic third graders performed at roughly the 40th percentile. Rather than narrowing over time, these early inequalities grew somewhat larger by eighth grade. Although initially lower scoring in ELA, Asian students caught up to their white peers by eighth grade. Our findings are similar in math, except that Asian

[4] The high school analyses are taken from larger, stand-alone studies recently completed or under completion by the authors. For further information about these studies, please contact the first author at ddr2111@columbia.edu.

[5] Research using nationally representative data suggests that sizable racial/ethnic differences in standardized test scores are present on the first day of kindergarten (see Lee & Burkam, 2003). As currently organized, NYC public schools clearly have limited capacity to address inequalities that formed prior to students' entry into the school system. What the NYC public schools do with students after the start of formal schooling is of course the focus of this series of commissioned papers.

students exhibited an initial mathematics advantage over all other racial/ethnic groups that increased during elementary and middle school. The consistently lower average academic performance of black and Hispanic students leading up to high school has obvious implications for subsequent inequalities.

Student Mobility, Attendance, and Behavior

The average student in this cohort attended roughly three schools between kindergarten and the end of middle school.[6] However, black and Hispanic students experienced an above-average number of school transitions. Conversely, whites experienced fewer school changes than the typical student, Asian students experienced almost one less school transition than average, and Limited English proficient (LEP) students attended roughly 1.5 fewer schools. As can be said about many student subgroups, LEP students clearly do not represent a demographically homogenous group. Student mobility is a case in point: the average Hispanic LEP student attended one-half more K-8 schools compared to the average Asian LEP student.

In terms of school attendance, the typical student missed almost 6 days of school in sixth grade, 6.5 days in seventh grade, and almost 8 days in eighth grade. Students with individualized education programs (hereafter referred to as IEP students) missed 75% more days of school than non-IEP students per year, while black and Hispanic students had above-average absence rates. Across all racial/ethnic groups, males missed more days of school compared to females. Finally, 6.7% of students were suspended at least once in middle school. However, over 15% of students with an IEP received an out of class suspension, as did over eight percent of black students—almost twice the rate of whites and almost four times the rate of Asian students. Student academic achievement was also associated with suspension; the proportion of low-achieving students experiencing middle-school suspensions was double that of their higher-achieving peers.

School Academic and Demographic Compositions

Students in this cohort attended elementary and middle schools that were racially and ethnically quite segregated. In kindergarten, the school attended by the average black student had

[6] School moves include planned moves between schools with different grade configurations, unplanned moves within NYC public schools, moves to private schools in NYC, and moves outside of NYC.

a 65.7% black enrollment, while the average Hispanic student attended kindergarten in a school that was 60.5% Hispanic. That year only 16% of kindergarteners citywide were white, but the average white kindergartener in this cohort attended a school that was 51.7% white. However, as this cohort progressed from third grade to middle school to high school, their schools became somewhat less segregated as school membership became less tied to neighborhood residence.

While schools became slightly less racially and ethnically segregated as this cohort progressed through the system, they also became more academically stratified. This stratification was particularly pronounced at the high school level in mathematics. However, counter to a common perception, the majority of test-score variability across all grades and subjects was within (rather than between) schools. In other words, test-score differences between two students randomly selected in any given school were much more likely to differ than the school-average means of any two randomly selected schools.

School-to-School Pathways

The final set of K-8 analyses explored whether certain students consistently attended schools serving higher-needs student populations.[7] Again, in contrast to public perception, there was considerable shuffling across school types over time. However, early school placements were clearly associated with subsequent school assignments. Over half of students who attended higher-needs schools in third grade went on to attend higher-needs schools for seventh grade, compared to roughly 15% of students who attended a lower-needs school in third grade. Importantly, one in three black and Hispanic students attended at least one higher-needs school in either third, seventh or ninth grade; among Asian and white students it was one in ten. Approximately three out of four students who never attended a higher-needs school graduated within four years. Of those who attended higher-needs schools in all three grades only one in three graduated within four years.

[7] We created a composite measure that includes each school's grade-specific ELA and math test scores (scale scores, reverse coded), mobility rate, attendance rate, percent of students receiving free/reduced-price lunch, and percent of students with an IEP. We then identified schools in the top third of this measure's distribution, which we labeled "higher-needs schools." It is important to bear in mind that this cut point is quite arbitrary, and only indicates need in a relative sense, as many "lower-needs" schools still face challenges only slightly less severe than those faced by schools designated "higher-needs."

PART 2: Student Outcomes and Experiences in High School

The results described above indicate the dramatically disparate academic skills that this student cohort brought with them to high school. Paths toward (and away from) college and career readiness were clearly being paved long before ninth grade. In the analyses that follow we describe what happened to these students as they progressed through high school and--for a select few--what happened when they entered college. We begin by describing student characteristics associated with on-time (four-year) high school graduation. From these findings we identify key factors associated with graduation, paying particular attention to student characteristics that are arguably within a high school's ability to influence: ninth-grade credit accumulation, ninth-grade attendance, and out of class suspensions. The next set of analyses explores students' experiences within the mathematics curriculum, which are closely linked to post-secondary success.

High School Graduation

Graduation rates in New York City have increased considerably over the past decade.[8] Indeed, these improvements are among the most substantial ever achieved by a large U.S. urban school district. Despite these laudable advances, the fact remains that only two-thirds (67.1%) of students in this cohort graduated within four years (by Spring, 2009).[9] Figure 1 displays the student-level attributes associated with four-year graduation, holding constant the other student-level traits. Remarkably, even after accounting for a host of other student characteristics, the odds of graduation for a student who completed fewer than 10 ninth-grade course credits (five year-long courses) were 85% lower, while students who missed more than 15 days of school in ninth grade had odds of graduating that were over 60% lower.[10] On an unadjusted basis, only

[8] See Kemple (2013) for more detailed information about recent trends in high school graduation rates, as well as a helpful overview of the current state of NYC high schools.

[9] Our estimated graduation rate is actually somewhat higher than the rate calculated by the NYC DOE. This is largely due to the fact that our sample does not include students not enrolled in a NYC DOE school in ninth grade, nor does it include those who were enrolled in specialized programs serving students in the criminal justice system, drug/alcohol rehabilitation centers, teen pregnancy centers, mental health facilities, etc., very few of whom graduated within four years. We also conducted analyses that considered graduation in five years. This raised the graduation rate to roughly 71%. Using the five-year graduation outcome did not demonstrably alter the results reported here.

[10] Graduation requires 44 credits, or 11 credits per year.

27.9% of students who earned less than 10 credits in ninth grade graduated, as did only 22.5% of students with 15 or more absences in ninth grade. These findings lend clear support to recent national efforts in large urban districts to identify struggling students early in their high school careers. Students who were ever suspended in high school and those who changed schools during high school were also considerably less likely to graduate on time.[11] Similarly, IEP, Hispanic, older, and LEP students all had odds of graduating that were between 13-33% lower.[12] Conversely, females were 25% more likely to graduate compared to their male peers, all else equal.[13] Unsurprisingly, students with higher levels of prior achievement were more likely to graduate within four years.[14] It is important to note that all else equal, black and white students had statistically similar odds of graduating.[15] Unfortunately, as we described above in the K-8 analyses, all things were clearly *not* equal.

[11] 21.6% of students in this cohort changed schools during high school.
[12] Each one-year increase in age was associated with a roughly 20% reduction in the odds of graduation.
[13] We also tested race/ethnicity by gender interactions. None were significant in the fully adjusted models.
[14] Specifically, each one-standard deviation increase in eighth-grade ELA and math tests scores was associated with a 39.1 and 62.8% increase in the odds of graduating within four years, respectively.
[15] The FRPL estimate was also non-significant, but we put little weight on this measure in this study given the tremendous noise inherent in the variable. According to the NYC Independent Budget Office (2011), approximately 17% of students do not return the lunch forms or return incomplete forms, and therefore their eligibility remains unknown.

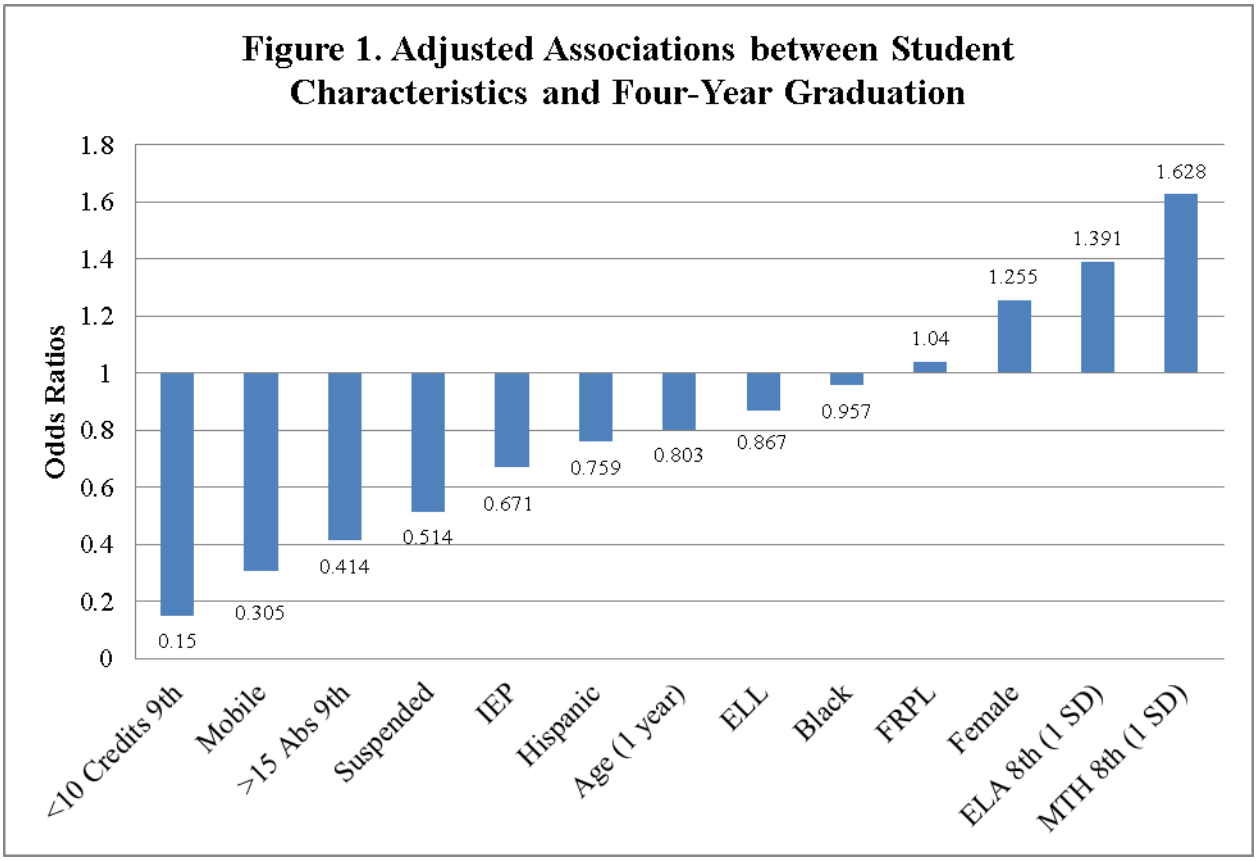


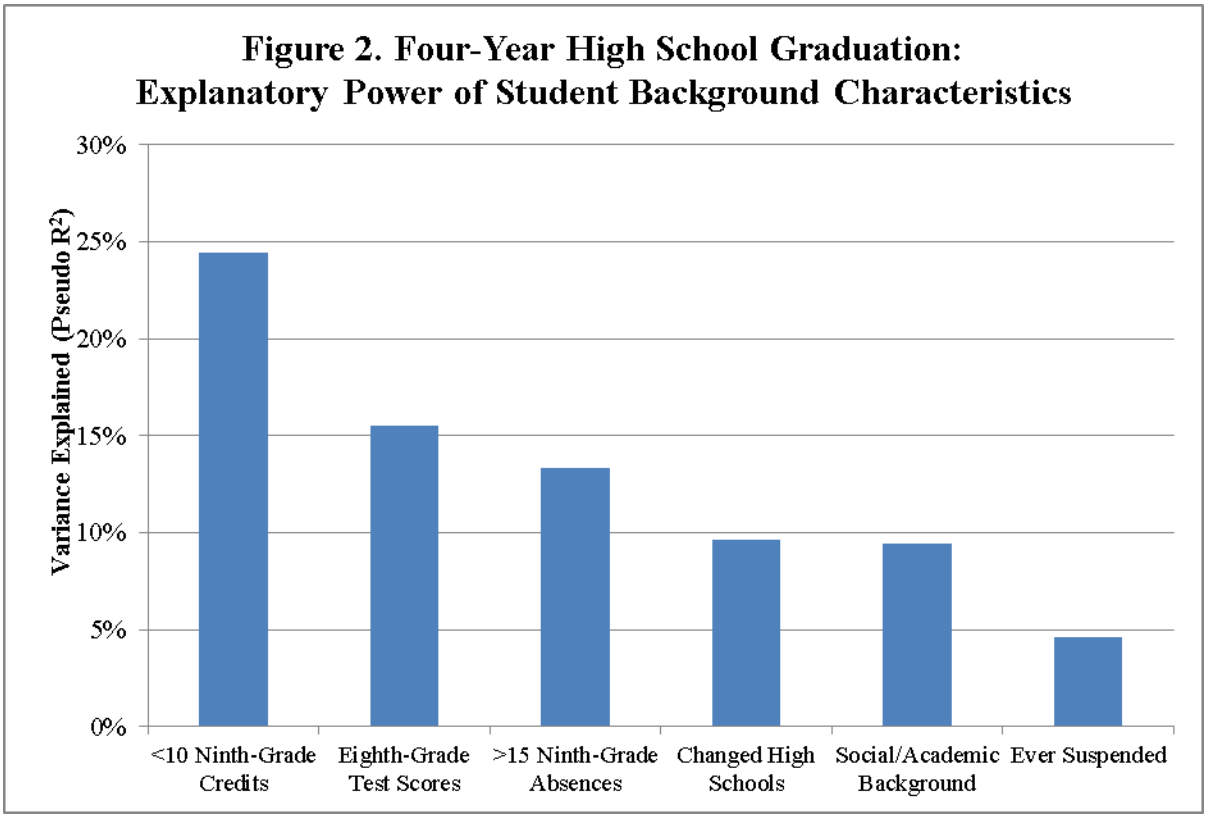
Figure 2 indicates the student-level attributes (or groups of attributes) most strongly associated with high school graduation.[16] Our findings closely mirror those reported by studies using data from New York City as well as other large urban districts.[17] Almost one-quarter of the variance in graduation is explained solely by whether students accumulated 10 or more credits in ninth grade, while eighth-grade ELA and math test scores together explain roughly 15%. Ninth-grade absences and high school mobility were third and fourth in terms of their ties to four-year graduation rates. Race/ethnicity, gender, IEP and LEP status, age, and free/reduced-price lunch (FRPL) status together explained less than 10% of the variance in four-year graduation.[18]

[16] Following Allensworth and Easton (2007), these values represent the reduction in log-likelihood (pseudo-R^2) when each of the six measures (or sets of measures) are included separately in six separate logistic regression models. However, because pseudo-R^2s associated with logistic regression are not directly comparable to R^2s produced by traditional OLS regressions, we include them here only as a rough guide. They should not be interpreted as precise estimates of variance explained.

[17] See Kieffer, Marinell, & Stephenson (2011) for similar findings from NYC, and Allensworth & Easton (2007) for similar findings from Chicago.

[18] Although not altogether appropriate given the dichotomous outcome, we also employed a traditional OLS regression (a linear probability model) to gain further insight into the relative strength of the links

Whether students were ever suspended during high school explained roughly 5% of the variability in the likelihood of graduating. While two of three students in this cohort graduated in four years, only four in ten suspended students did so. We now focus on three of these six indicators of high school graduation. Unlike eighth-grade test scores, student mobility, or student social and academic backgrounds, high schools can potentially influence ninth-grade attendance rates and credit accumulation, and address behaviors and implement policies to reduce suspensions.



Ninth-Grade Credit Accumulation

The extant research also points to the importance of ninth-grade credit accumulation in keeping students on track to graduate.[19] Model 1 in Table 1 indicates that black students in this cohort finished ninth grade with 1.35 fewer credits compared to white students, while Hispanic

---

between student characteristics and the likelihood of graduation. In fully-adjusted models, roughly one-third of the explained variance in graduation rates was credited to ninth-grade credit accumulation. Indeed, ninth-grade credit accumulation told us more about the likelihood of graduation (from a statistically adjusted standpoint) than race/ethnicity, age, IEP, LEP, and FRPL status *combined*.

[19] See again Allensworth & Easton (2007).

and other race/ethnicity students ended the year roughly 1 credit behind.[20] Conversely, Asian students finished at a 0.33 credit advantage over their white peers. Model 2 suggests that all else equal, females gained almost 0.75 more credits, LEP students earned roughly 0.5 more credits, and each additional year of age was associated with roughly 0.5 fewer credits completed in ninth grade. By far the strongest predictor here was IEP status: special education students passed almost 2 fewer credits.[21]

In Model 3 we (unsurprisingly) find that students with higher eighth-grade test scores earned more credits in ninth grade. A key set of findings here relate to students' opportunities to learn. Specifically, students who experienced a ninth-grade suspension completed one fewer credit. Similarly, all else equal, an increase of ten absences is associated with the completion of one fewer credit, while an increase of ten late arrivals is associated with 0.3 fewer credits earned. Students who changed schools during ninth grade also earned almost one-half fewer credits. As we discuss below, changing schools is associated with several weeks of additional school days missed. Based on data from the NYC DOE annual school surveys, Model 4 indicates that students attending schools where they felt safe earned more ninth-grade credits.[22] We also find support here for the New York City small high schools. When we account for the fact that they typically enrolled larger proportions of socially and academically disadvantaged students, our results suggest that students attending small, non-selective schools earned more credits in ninth-grade compared to students who attended most other types of schools.[23]

[20] Compared to ninth-grade absences, ninth-grade credit accumulation varied somewhat more across schools. Specifically, 76% of the variability in the number of credits ninth-graders in this cohort completed existed within schools, while 24% was found between schools.

[21] Note, however, that the IEP estimate is rendered non-significant in subsequent models (due largely to the inclusion of eighth-grade test scores).

[22] A one standard-deviation increase in members' views of school safety was tied to passing roughly one-half more courses.

[23] In unadjusted models, we still find an advantage of small, non-selective schools over medium/large non-selective schools, whose students earned 1.76 fewer credits during ninth grade. District 75 and transfer-school students also earned far fewer credits. There were no unadjusted differences between other school types and small, non-selective schools. Bear in mind, however, that the outcome here does not consider the rigor of the credits earned.

Table 1. Ninth-Grade Credits Earned (*n*=70,130)

	Model 1	Model 2	Model 3	Model 4
Student Characteristics				
Black[1]	-1.35***	-1.27***	-0.54***	-0.55***
Hispanic	-1.01***	-1.04***	-0.37***	-0.37***
Asian	0.38***	0.27***	-0.18*	-0.18*
Other Race/Ethnicity	-0.99***	-0.90***	-0.66**	-0.67**
Female		0.71***	0.67***	0.67***
IEP		-1.86***	0.14	0.14
LEP		0.46***	0.03	0.03
Free/Reduced Lunch		-0.03	-0.04	-0.04
Age (years)		-0.53***	-0.08**	-0.08**
8th Grade Math Ach.[2]			0.75***	0.75***
8th Grade ELA Ach.[2]			0.28***	0.28***
Suspended in 9th Gr.			-1.08***	-1.08***
Days Absent 9th Gr.			-0.10***	-0.10***
Days Late 9th Gr.			-0.03***	-0.03***
Changed Schools 9th			-0.43**	-0.43**
School Characteristics				
Small, Selective[3]				-1.04***
Med/Large Non-Selective				-1.69**
Med, Selective				-0.74*
Large, Selective				-0.98*
District 75				1.73
Transfer School				-3.19***
Academic Expectations[2]				-0.32
Safety[2]				0.55*
Communications[2]				-0.21
Engagement[2]				0.42
Avg. 8th Grade Math Ach.[2]				-0.37
Avg. 8th Grade ELA Ach.[2]				0.19
% IEP				-0.03
% LEP				-0.02*
% FRPL				0.00
% Black				0.01
% Hispanic				0.02*
Intercept	12.06***	11.90***	11.46***	12.17***

~*p*<.10; *p<.05; **p<.01; ***p<.001 using robust standard errors.
[1] All racial/ethnic groups compared to white students;
[2] Measure is standardized (z-scored; M=0, SD=1)
[3] All school classifications compared to small, non-selective schools;

Ninth-Grade Absences

As noted above, ninth-grade absences were a critical predictor of on-time graduation. On average, students in this cohort missed over three weeks (17 days) of school in ninth grade.[24] Table 2 presents analyses that explore the extent to which school attendance rates during this crucial year varied by student and school characteristics.[25] Model 1 indicates the unadjusted racial/ethnic differences in ninth-grade student absences, which were not particularly large. The average white student missed between 16 and 17 days of school (as indicated by the intercept). The typical black or Hispanic student missed roughly two more school days than this in ninth grade, while Asian students on average missed almost four fewer days. Models 2 and 3 adjust these racial/ethnic differences for additional student characteristics. Portions of the black/white and Asian/white attendance disparities are explained by these other student attributes. This is largely because special education status, which is more prevalent among black students, was associated with eight additional absences, while LEP status (more prevalent among Asian students) was associated with a four-day reduction in absences. Older students also missed more days that year, with each additional year in age associated with a three-day increase in the number of days missed.

Model 3 suggests all else equal, black students actually missed fewer days than predicted, given their relatively more disadvantaged backgrounds. Note that students who were suspended at any point in ninth grade missed over two weeks of additional school, even after controlling for their other characteristics. Importantly, as we discuss below, black students were far more likely than white students to be suspended in ninth grade. These absences associated with out of class suspensions are above and beyond those associated with the suspensions themselves, as suspended students in New York City are only considered absent if they do not report to their alternate assigned school or instructional time. Students who changed schools in ninth grade also missed over three weeks more school. Higher levels of prior achievement, particularly eighth-grade mathematics test scores, were also associated with fewer absences.

[24] The average attendance rate, however, masks the fact that a relatively small number of chronically absent students pulled the average absence rate upwards. The median student in this cohort missed 10 days of school in ninth grade.

[25] Importantly, among students in this cohort, variability in ninth-grade school absences was largely a within-school phenomenon. Over 85% of the variability in student absences was found to exist within schools (and thus only 15% between schools). This suggests that the extent to which student attendance rates can be explained as a function of school characteristics is somewhat limited.

In terms of school characteristics, our results again suggest positive outcomes associated with New York City's small high schools.[26] Students in this cohort who attended medium-sized and large non-selective high schools missed almost four more days in ninth grade compared to students attending small, non-selective high schools, all else equal.[27] Schools in which members felt a greater sense of academic expectations had lower absence rates, with a one-standard-deviation increase in academic expectations associated with an almost two-day average absence reduction. Higher concentrations of black and Hispanic students were also associated with increased absences, even after controlling for student-level race/ethnicity.

[26] These descriptive findings regarding NYC's small schools closely mirror the positive findings reported by Bloom, Thompson, & Unterman (2010) and Bloom & Unterman (2012). We strongly recommend these papers to interested readers.

[27] However, we again caution readers that despite the rich controls employed here, these remain associational rather than causal relationships.

Table 2. Ninth-Grade Student Absences (*n*=70,130)

	Model 1	Model 2	Model 3	Model 4
Student Characteristics				
Black[1]	1.83***	1.18***	-0.87*	-1.05**
Hispanic	2.54***	2.69***	1.57***	1.47***
Asian	-3.62***	-2.95***	-2.01***	-2.04***
Other Race/Ethnicity	1.28	0.62	-1.22	-1.34
Female		-0.27	0.00	0.03
IEP		7.77***	3.68***	3.56***
LEP		-4.29***	-5.28***	-5.18***
Free/Reduced Lunch		0.10	0.36*	0.34*
Age (years)		3.32***	2.59***	2.60***
8th Grade Math Ach.[2]			-3.16***	-3.10***
8th Grade ELA Ach.[2]			-0.27*	-0.21
Suspended in 9th Gr.			10.15***	10.11***
Changed Schools 9th			15.78***	15.59***
School Characteristics				
Small, Selective[3]				-0.56
Med/Large Non-Selective				3.84***
Med, Selective				-0.87
Large, Selective				0.26
District 75				8.18
Transfer School				1.44
Academic Expectations[2]				-1.80~
Safety[2]				-0.60
Communications[2]				0.94
Engagement[2]				0.16
Avg. 8th Grade Math Ach.[2]				-0.78
Avg. 8th Grade ELA Ach.[2]				1.42
% IEP				0.10
% LEP				0.00
% FRPL				0.03
% Black				0.09***
% Hispanic				0.06**
Intercept	16.65***	16.31***	16.81***	16.45***

~*p*<.10; **p*<.05; ***p*<.01; ****p*<.001 using robust standard errors.
[1] All racial/ethnic groups compared to white students;
[2] Measure is standardized (z-scored; M=0, SD=1)
[3] All school classifications compared to small, non-selective schools;

High School Suspension

Discussions about how to ensure high school graduates are career and college ready typically—and understandably—focus on curricula, instruction, and school organization. However, a crucial piece often neglected is that students who are not in school are unable to benefit from improved curricula, better instruction or well-designed and well-organized schools. As we reported above, students in this cohort who were suspended, absent, late, and mobile were less likely to be successful in ninth grade. In this section we focus on out of class suspensions during high school. These suspensions were disproportionately experienced by particular student populations, and, as we explore in a later section, were tightly linked with other negative outcomes.

It is important to stress that the analyses below simply describe *which students* in this cohort were most likely to experience a suspension. They do not (and cannot) address the fairness or objectivity of the suspensions these students received. We should also note that since 2010, the NYC DOE and the New York City Council have taken measures to revise the discipline code and school safety policy to bring more transparency to the system and to decrease the number and length of suspensions. In 2010, the Council passed the Student Safety Act, which required that school security officers report arrests and summons executed at schools and that schools report detailed suspension and expulsion data. That year, the NYC DOE also decreased the number of zero tolerance infractions listed in the discipline code from 29 to 21. In 2012, the NYC DOE took additional measures to limit the number and length of suspensions, including capping the number of days young children can be suspended for low to mid-level offences. The students in this cohort, however, were obviously not able to benefit from any positive outcomes these policy changes may have brought about. The findings we report below may or may not have differed under the recent policy changes.

Descriptive differences. Table 3 displays academic and socio-demographic differences between students who were and were not suspended at least once in high school. Roughly 15% of the sample—over 11,000 students—had been suspended. These descriptive results generally confirm prior research regarding the types of students more likely to be suspended. Students who were suspended were almost twice as likely to have IEPs, and male suspended students outnumbered females almost two to one. In terms of racial/ethnic background, the clear finding is the overrepresentation of black students among those suspended. Black students constituted

less than one-third of the cohort but approximately half of all suspended students. Conversely, white and Asian students were far less likely to be suspended. We also find that those suspended were less likely to be LEP students, but were typically somewhat older than students who were not suspended. School suspensions were also associated with much weaker academic backgrounds, with suspended students typically entering high school with lower eighth-grade standardized test scores in both ELA (-0.411 *SD*s) and mathematics (-0.472 *SD*s).

During high school, suspended students missed over seven more days of school per semester, and were late over four more days. Suspended students also typically passed a much smaller proportion of their classes. Whereas students who were never suspended passed over 80 percent of their classes, suspended students passed fewer than 60 percent. As a result, they earned roughly 1.6 fewer credits per semester. Broken down by subject, suspended students gained approximately 30 percent fewer credits in math and 25 percent fewer credits in English per semester compared to their peers who were never suspended.[28] Perhaps most importantly, only 40 percent of students who received a suspension graduated within four years, compared to 72 percent of their non-suspended peers.

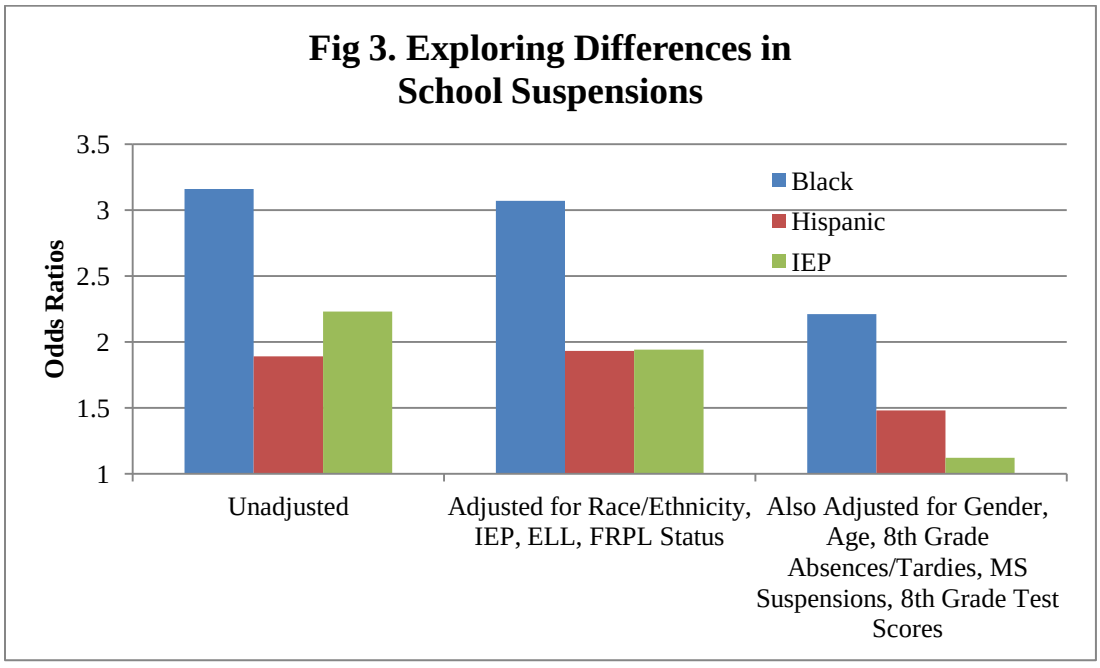
[28] Analyses not presented here included student fixed-effects models that compared credits earned and absences in semesters when students were suspended compared to semesters in which they were not suspended. One benefit of this approach is that it removes unmeasured differences across students associated with both the likelihood of being suspended and credit accumulation and attendance. Although the effects were much smaller, we still found that students earned fewer credits and had more absences in the semesters in which they were suspended.

Table 3. Associations between School Suspension and Students' Socio-Demographic and Academic Characteristics (*n*=70,130)

	No Suspensions (n=58,967)	At Least One Suspension (n=11,163)
Background Characteristics		
% Special Education	9.6***	19.1
% Free/Reduced-Price Lunch	45.8	44.8
% Female	52.9***	37.8
Race/Ethnicity***		
% White, non-Hispanic	14.2	7.3
% Black, non-Hispanic	31.1	50.8
% Hispanic	37.3	36.6
% Asian	16.7	4.7
% Other	0.7	0.6
% Limited English Proficiency	16.1***	9.6
Age at 9th Grade Entry	14.4***	14.5
Middle School Academic Outcomes		
Eighth Grade ELA Achievement	0.029***	-0.382
Eighth Grade Math Achievement	0.067***	-0.405
Eighth Grade Absences	14.1***	19.4
Eighth Grade Lateness	14.0***	24.6
% Received Suspension in Middle School	4.0***	15.2
High School Academic Outcomes		
Days Absent Per Semester	10.1***	17.5
Days Late Per Semester	5.9***	10.8
% Credits Passed Per Semester	81.4***	59.7
Credits Earned Per Semester	6.0***	4.4
% Math Credits Passed Per Semester	77.2***	54.8
Math Credits Earned Per Semester	0.9***	0.7
% ELA Credits Passed Per Semester	82.6***	60.7
ELA Credits Passed Per Semester	1.2***	0.9
% Graduated in 4 Years	72.2***	40.2

***$p<.001$

Unpacking suspensions. The next set of analyses disentangles the complicated associations between student characteristics and the likelihood of school suspension. We focus on three groups of students who experienced a disproportionate number of suspensions and were also less likely to graduate from high school: students with an IEP, black students, and Hispanic students. The first set of bars in Figure 3 presents the unadjusted IEP/non-IEP, black/white and Hispanic/white suspension odds ratios. An odds ratio of 1.0 indicates that two groups have equal odds of suspension, while odds ratios above 1.0 indicate greater odds, and odds ratios below 1.0 indicate lesser odds of being suspended. Reflecting the descriptive findings in Table 3, we find that, compared to white students, Hispanic students were twice as likely to be suspended and black students three times as likely. IEP students had odds of being suspended that were twice those of non-IEP students. The second set of bars in Figure 3 is adjusted for race/ethnicity, and IEP, LEP, and free/reduced-price lunch status. Doing so, however, explains very little of the differences in suspension rates among groups. The final set of bars indicates the results from the full model including all student-level characteristics. Even after adjusting for a host of student social and academic background characteristics, we still find that black students in this cohort were over twice as likely to be suspended compared to whites, the odds that a Hispanic student would be suspended were over 50 percent greater, and the odds that a student with an IEP was suspended were roughly 10 percent greater.



Quasi-experimental approach. We also used propensity score matching to estimate the effect of earning at least one suspension on the odds of graduating from high school in four years. This technique allows us to match suspended students to non-suspended students based on observable characteristics, thereby restricting our sample to similar groups of students. Here, we used a multi-level matching approach, with students nested within their ninth grade school and matched on eighth-grade characteristics including test scores, attendance, and suspension history, as well as socio-demographic characteristics.[29] Our results suggest that earning a suspension was associated with a 46 percent decrease in the odds that students graduated from high school in four years. In other words, had these suspended students never earned a suspension, their odds of graduating would have been almost twice as great.

High School Mathematics Coursetaking

Our final set of high school analyses focuses on mathematics coursetaking, which is closely linked to post-secondary educational success. In particular, progressing to at least Algebra II seems to be a critical determinant of college enrollment and completion.[30] Unfortunately, low-income and minority students typically complete fewer advanced mathematics courses during high school.[31] As do the graduates of most large, urban school districts, students in this New York City cohort struggled with mathematics in their post-secondary schooling. Among students who entered the CUNY system in Fall 2009—the semester these graduates would have begun had they enrolled immediately after high school graduation—little more than half (52.2%) passed a credit-bearing math course within the first two years of study.[32]

The analyses below focus on two simple questions: Which students took which types of mathematics courses, and to what extent did the rigor or expectations associated with those courses vary across student and school characteristics? Instead of simply calculating the number of math credits taken by students, we measure student coursetaking using the highest level of math passed during high school to assess how far students progressed through the math curricular pipeline (e.g., 0=no math, 8=advanced topics beyond calculus; see Table 4). This

[29] See Appendix A for a discussion of the propensity score approach.
[30] See Adelman (2006).
[31] See Lee & Ready (2009) and ACT (2006).
[32] See CUNY (2013).

technique emphasizes *which* courses students took during high school over how many years of courses they took, and is more consistent with the literature on what matters both for learning during high school and for college and career readiness. With the pipeline measure one unit approximates one year of math coursetaking and 0.5 units approximates one semester. AP math courses are separated into their own category, although few students took both the regular and AP version of the same course. However, AP calculus, for example, covers more content than regular calculus. Moreover, AP courses serve as important signals to colleges regarding students' math ability and preparation. For these reasons, we consider reaching AP/IB math courses as progressing further through the math curriculum compared to the regular versions of those same courses.

Table 4. Curriculum Equivalents of the Math Pipeline Measure

Math Pipeline Level	Curriculum Equivalent
0	No math
1.0	Pre-Algebra, Basic math, Consumer math, etc.
1.5	Algebra term 1of 2, Math A term 1 of 3
2.0	Algebra term 2 of 2, Math A term 2 of 3
2.5	Geometry term 1 of 2, Math A term 3 of 3
3.0	Geometry term 2 of 2, Math B term 1 of 3
3.5	Algebra II/Trigonometry term 1 of 2, Math B term 2 of 3
4.0	Algebra II/Trigonometry term 2 of 2, Math B term 3 of 3
4.5	Pre-Calculus term 1 of 2
5.0	Pre-Calculus term 2 of 2
5.5	Calculus term 1 of 2
6.0	Calculus term 2 of 2
6.5	AP/IB Calculus or AP/IB Statistics term 1 of 2
7.0	AP/IB Calculus or AP/IB Statistics term 2 of 2
8.0	Advanced topics (e.g., Multivariate Calculus)

Who takes what? Figure 4 displays (unadjusted) average progress through the math pipeline for students differentiated by their background characteristics. We find substantial racial/ethnic disparities in math coursetaking. The average Asian student completed the first semester of pre-calculus and the average white student progressed through both semesters of algebra II/trigonometry. By contrast, the average black or Hispanic student only completed the second semester of geometry. The NYC DOE considers only math courses that are at the level of algebra II or beyond to be "college preparatory" when giving schools credit for promoting

college and career readiness (NYC DOE, 2011). Thus, by this definition the average black or Hispanic student in this cohort completed no college preparatory courses.

Also evident in Figure 4, LEP and free/reduced-price lunch students appear to have progressed less far in math, on average. Substantively, however, neither of these disparities are particularly large. On the other hand, we find a large disparity in math coursetaking between students with IEPs and those without. Students with IEPs, on average, progress only through the first semester of geometry, while non-IEP students progressed through the first semester of algebra II/trigonometry. Finally, we find substantial differences in progress through the math pipeline by students' initial achievement level (as measured by eighth grade math test scores). Low-achieving students progressed, on average, only through the first semester of geometry, while average-achieving students progressed through the first semester of algebra II/trigonometry, and high-achieving students advanced all of the way through pre-calculus.[33]

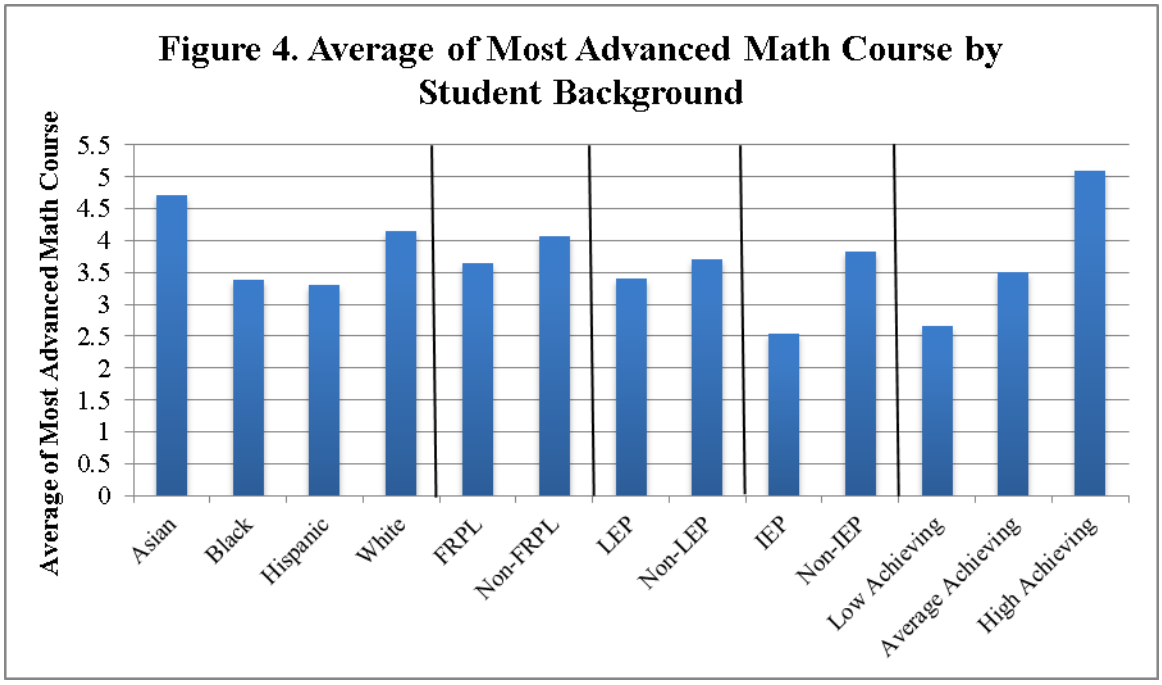


Table 5 displays the results of our analyses of the student and school characteristics associated with greatest progression through the mathematics pipeline. Unlike the results on

[33] Low achievement is defined as eighth-grade math test scores below -0.5 SDs, high achievement scores are above 0.5 SDs, and average scores fall between -0.5 and 0.5 SDs.

coursetaking presented thus far, here we adjust for students' socio-demographic and academic characteristics. This allows us to assess, for example, whether racial/ethnic gaps in mathematics coursetaking remain, even after controlling for attributes such as students' initial math achievement, poverty, and LEP status. During their tenure in high school, Asian students advanced roughly two-fifths of a year further in math than their white counterparts, all else equal. In contrast, both black and Hispanic students lagged behind their white peers by nearly half a year in math coursetaking. Native American/Multi-ethnic students progressed approximately one-quarter of a year less far through the math pipeline than white students.

Female students advanced nearly one-sixth of a year further than males in math, suggesting that the traditional gender gap favoring males had reversed for this cohort. IEP and LEP status were both negatively associated with progress in math. FRPL eligibility was surprisingly slightly positive and significant in Model 1, although it becomes negative in later models, reinforcing the concern that this is a very weak measure of students' socioeconomic status. It is important to note that the constant in Model 1, which reflects the most advanced math course taken in four years by the *average* student, is 3.82. In other words, the average student in this cohort (excluding transfers and dropouts) progressed to roughly the first semester of algebra II/trigonometry.

In Model 2, we further adjust for students' behavioral and academic characteristics prior to entering high school. Predictably, entering math achievement was strongly related to coursetaking, with a one standard-deviation increase in math achievement on entering high school associated with one-semester further advancement through the math curriculum. After adjusting for math achievement as well as the other academic and behavioral attributes in Model 2, the racial/ethnic disparities in math coursetaking are reduced but do not disappear. Model 3 indicates that a one-year increase in the level of students' first math course attempted in high school is associated with over two-fifths of a year increase in eventual four-year progress through the math pipeline. Racial/ethnic disparities in advancement through the math curriculum are again slightly reduced after controlling for first math course but remain significant, indicating that they are not purely a function of where students were when they entered high school but what occurred during high school as well. Put another way, even after adjusting for where they started, Asian students progressed further in the math curriculum than white students, who in turn progressed further than black and Hispanic students.

Model 4 adjusts for school factors including school type, student body composition, curriculum structure measures, and aggregate student perceptions of school climate. Interestingly, after controlling for all other student and school characteristics, students attending small, non-selective schools appear to have progressed further through the math curriculum, on average, than students in all other school types. Specifically, students in the small non-selective schools progressed further through the math curriculum than students in the large and midsize nonselective schools by over one semester. We also find several significant associations between the attributes of schools' curriculum structure and students' progress in math. Students who attended schools that offered calculus advanced further through the math curriculum than students who attended schools that did not.[34] In contrast, we find a slight negative association between the percent of courses schools offered below Algebra and student progress in math. A 20 percent increase in the percent of courses below Algebra is associated with approximately one-seventh of a year less progression. Although the magnitude of the association is not substantial, this finding does suggest that eliminating lower level courses may have a positive influence on student coursetaking.[35] Finally, all else equal, a one standard deviation increase in school-members' sense of the level of academic expectations is associated with a nearly one-third year further advancement through the math curriculum.

[34] Although our data include thousands of College Now (CN) math courses, we cannot be completely sure that we have captured all of them, as not all schools during this period were required to report their students' CN coursework. We do, however, know that of the 48,700 2009 graduates, 19% took a CN course in any subject. Of those graduates, just under half (49.7%) passed Algebra II. Assuming for the moment that each and every CN course completed was an advanced mathematics course, an obviously unreasonable assumption, at most 9.5% of graduates could have theoretically completed an advanced CN math course beyond Algebra II.

[35] On a more technical finding, note the positive relationship between the within-school *SD* of the most advanced math course taken and students' progress in math. Viewed as a measure of within-school curricular differentiation, the positive coefficient indicates that, on average, students in more differentiated schools progressed slightly further through the pipeline than students in more constrained schools. It is possible that schools with narrow curriculums more often target their courses at a lower level of difficulty, artificially constraining the average student's progress in math to be lower than if allowed to freely manifest itself in a broader curriculum with more advanced course offerings. This is an instance where average coursetaking may be positively influenced, but that the equitable distribution of that coursetaking may be diminished. We plan to address this question in future research.

Table 5. Multi-Level Models of Highest Math Course Completed[a] (*n*=45,855)

	Model 1	**Model 2**	**Model 3**	**Model 4**[b]
Socio-Demographic Characteristics				
Asian[c]	0.42***	0.26***	0.25***	0.24***
Black	-0.47***	-0.23***	-0.18***	-0.18***
Hispanic	-0.47***	-0.30***	-0.25***	-0.25***
Native/Multi-Ethnic	-0.28***	-0.12	-0.08	-0.08
Female	0.16***	0.14***	0.14***	0.14***
Free/Reduced-Price Lunch	0.04*	-0.05***	-0.05***	-0.05***
Special Education	-0.77***	0.04*	0.01	0.01
English Language Learner	-0.19***	0.06**	0.06**	0.06**
Behavioral/Acad. Characteristics				
Overage for 8th grade		-0.03	-0.07***	-0.07***
Middle school mobility		-0.05***	-0.06***	-0.06***
8th grade math achievement[d]		0.52***	0.46***	0.46***
8th grade ELA achievement[d]		0.16***	0.11***	0.11***
Days absent in 8th grade		-0.02***	-0.02***	-0.02***
Days tardy in 8th grade		-0.01***	-0.01***	-0.01***
First Math Course[e]			0.57***	0.57***
School Characteristics				
School Type[f]				
Small, Selective				-0.12*
Midsize, Selective				-0.21**
Midsize & Large, Nonselective				-0.52***
Large, Selective				-0.36***
School offers Calculus				0.21**
% credited courses below Algebra				-0.01*
Within-school SD of most adv. course[d]				0.11***
Academic expectations[d]				0.32***
School safety[d]				-0.10
Communications[d]				-0.04
Student engagement[d]				-0.10
Intercept	3.82***	3.89***	3.87***	3.93***

*p<.05, **p<.01, ***p<.001

[a] The most advanced math course passed in four years of high school (0=No math, through 8=Advanced topics beyond Calculus)

[b] The following non-significant school-level predictors were removed in the interest of model parsimony: borough indicators, the percent of courses classified as prep/support, whether the school is higher-needs or offers AP/IB math, and the percent of students that were FRPL eligible, special education, LEP, or overage.

[c] All racial/ethnic groups are compared to white students.

[d] Measure is z-scored (M=0; SD=1)

[e] Level of first math course taken in high school (0=No math, through 8=Advanced topics beyond Calculus)

[f] School type categories are compared to small, nonselective schools.

Assessing the Grade/Exam Score Performance Differential

We have thus far assumed that course titles are a good indicator of students' mastery of math content. The analyses have also implicitly assumed that how far a student advances in math is an equally good indicator of students' math proficiency across all schools in New York City. In this final section we assess the associations between the grades awarded students in their mathematics courses and their performance on the related New York State Regents exam. Specifically, we subtract the average mark students received in the courses (conditional on passing) covered by each Regents exam from their score on that Regents exam. This differential potentially reflects several different phenomena. It may indicate that certain students or those in particular schools are held to lower standards by their teachers. Alternately, substantially higher grades compared to exam scores may suggest that the course content of even similarly titled courses within (or between) particular schools is less advanced. Course grades also encompass more than content mastery, they also reflect persistence, effort, hard work and other extra-cognitive skills, including many that are directly relevant to college readiness. Because we cannot identify the specific mechanism at play, we agnostically term the resulting difference score the "performance differential," which ascribes no particular cause to the discrepancy. Clearly, course marks and Regents exam scores are not directly comparable metrics. As such, our focus here is how the performance differential *itself* varies across students and schools in a relative (rather than absolute) sense.[36]

[36] We examine the curricular rigor for course sequences leading to both the Math A and Math B Regents exams. From an analytic standpoint, both exams have strengths and weaknesses. Many more students took the Math A exam, covering roughly Algebra I and the first semester of Geometry, because it was required for graduation. Consequently the sample of students with Math A Regents exam scores is more representative of the overall NYC student population. On the other hand, because the content covered by the Math A exam is quite basic, many of the more advanced students take the first, and often second, courses of the Math A sequence in middle school. Because high schools can only control their own courses, we only include courses taken during high school. This suggests that for high-achieving students, the average course mark may only capture the last course or last two courses in their Math A sequence.

Conversely, the Math B Regents exam, which covers roughly the second semester of Geometry as well as Algebra II/Trig, was required only for students who wanted an Advanced Regents Diploma. As a result, many fewer students attempted the Math B Regents, and those students were typically more advanced. However, one advantage of the Math B exam for these analyses is that very few students begin the Math B course sequence in middle school, meaning that students' learning of Math B curricular content can be more fully attributed to their high schools.

Table 6 presents the relationships between student and school characteristics and students' performance differentials in Math A and Math B. We find that all else equal, Asian, black, and Hispanic students all had larger negative performance differentials in Math A than white students. In other words, white students' grades in their Math A courses were better indicators of how they performed on the less-subjective Math A Regents exams. Black and Hispanic students also had larger negative performance differentials in Math B than white students. Exploring the potential causes of these racial/ethnic disparities is crucial, but beyond the scope of this paper.

Female students were also generally awarded better grades than their test scores would suggest. IEP students had larger negative performance differentials in both Math A and Math B, although the discrepancy is much larger in Math A, most likely because few special education students with substantial learning disabilities took the Math B Regents. LEP and overage students had significantly larger negative differentials in Math A but not Math B. By contrast, students' eighth-grade math and English achievement was associated with smaller gaps between their grades and Regents scores.

Adjusted for student- and other school-level characteristics, the within-school mean of students' most advanced math course was *negatively* related to students' performance differentials in Math B, while the within-school standard deviation of students' most advanced math course—a measure of curricular differentiation—was *positively* related to their performance differentials in Math B. These results suggest that there may be a tension between getting more students through advanced math courses and maintaining the performance standards within those courses. In other words, after adjusting for characteristics such as student- and school-level entering math achievement, those schools with higher average coursetaking and less variation in coursetaking, actually had *worse* average performance differentials. We also find that schools that offer Advanced Placement or International Baccalaureate math courses had better average performance differentials in Math B than schools that do not. It is unlikely that offering an AP calculus or statistics course caused the rigor of Math B courses to go up, however it may reflect an underlying curricular mission of offering rigorous, college preparatory math courses.

In terms of school racial/ethnic composition, schools with greater concentrations of black and Hispanic students had larger discrepancies between student grades and their Math B exam

scores. Our results also suggest that school-average eighth-grade math achievement was related to students' performance differentials in both subjects, with higher achievement associated with stronger correlations between grades and exam scores. This may be because expectations for student performance were somewhat elevated in schools with higher average achievement. These findings regarding the percent of black and Hispanic students and average math ability are consistent with the extant literature, which finds that schools enrolling larger proportions of socially and academically disadvantaged students are more likely to award grades that are higher than standardized tests would suggest. Finally, attending a "higher-needs" school was associated with larger negative performance differentials in both Math A and Math B.

Finally, we investigated the relationships between school members' perceptions of school climate and their performance differentials. Importantly, in schools with increased academic expectations, student grades were more closely aligned with standardized measures of achievement.[37] We also find a positive association between school safety and students' performance differentials in Math A but not Math B. One finding that is more difficult to explain is the larger negative student performance differentials in Math B in schools with increased communications. One potential explanation is that at the high school level, where levels of communication are typically lower than in earlier grades, communication to students and parents is often of a negative nature, conveying information about academic or disciplinary problems. Once we adjust for the other student and school characteristics, student engagement was not significantly related to these mathematics performance differentials.

[37] A one standard deviation increase in schools' positive academic expectations is associated with decreases in students' negative performance differentials in both Math A (2.09 points) and Math B (6.34 points).

Table 6. Between-school, Multi-Level Models Exploring Student Performance Differentials[a]

	Math A		Math B	
	Model 1	**Model 2**	**Model 3**	**Model 4**
Student Characteristics				
Asian[b]	-0.40*	-0.41*	0.34	0.36
Black	-0.63**	-0.55**	-2.08***	-1.92***
Hispanic	-0.78***	-0.71***	-1.94***	-1.82***
Native/Multi	-1.29	-1.24	-1.43	-1.36
Female	-2.30***	-2.30***	-2.02***	-2.02***
Special Education	-7.43***	-7.46***	-2.05*	-2.05*
Limited English proficient	-1.52***	-1.53***	0.42	0.40
Free/Reduced-Price Lunch	-0.02	0.01	-0.27	-0.22
Overage for 8th grade	-1.05***	-1.05***	0.42	0.44
Middle school mobility	0.02	0.03	0.06	0.07
Entering math achievement[c,]	3.70***	3.67***	4.45***	4.42***
Entering ELA achievement[c]	1.40***	1.37***	0.98***	0.96***
Days absent in 8th grade	0.02**	0.02**	-0.04*	-0.03*
Days tardy in 8th grade	0.03***	0.03***	0.01	0.01
School Characteristics:				
Within-school mean most adv. course[c]		-0.46		-2.21*
Within-school SD most adv. course[c]		0.53		1.99**
School Offers AP/IB Math		0.98		4.78**
Percent Black and Hispanic		-0.01		-0.09**
Within-school mean 8th Gr. math ach[c]		1.65**		1.41*
Higher-needs school		-1.53*		-4.55*
Academic expectations[c]		2.09*		6.34**
School safety[c]		1.11*		1.28
Communications[c]		-1.33		-5.96**
Student engagement[c]		-1.10		-0.15
Constant	-7.47***	-7.59***	-26.95***	-28.25***

*p<.05, **p<.01, ***p<.001. The following school-level predictors, non-significant in all models, were removed in the interest of model parsimony: borough indicators, the percent of courses offered below Algebra or classified as prep/support, whether the school offers Calculus, school type categories, and the percent of students that were FRPL eligible, special education, LEP, or overage.

[a] Performance differential refers to the gap between students' average course marks and their Regents Exam marks for the same subject. It is calculate by subtracting the average course mark from the Regents mark.

[b] All racial/ethnic groups are compared to whites.

[c] Measure is z-scored (M=0; SD=1)

College Attendance and One-Year Persistence

As noted above, roughly two-thirds of students in this New York City high school cohort received a high school diploma in four years. Our final set of analyses focus on the extent to which these high school graduates subsequently enrolled in college. As we were completing this paper, data for this cohort were not yet available on college graduation, which would have occurred in Spring 2013 for students graduating in four years, or Spring 2015 for students graduating in six years. We do, however, have data on college enrollment and early college persistence. Over two-thirds of high school graduates in this cohort enrolled in college for the 2009-2010 academic year. Recent national estimates suggest that 63 percent of first-time students at four-year public institutions graduate within six years.[38] If we naively assume that the high school graduates in this cohort who enrolled in college all did so as full-time students in four-year public institutions, a liberal estimate is that roughly 42 percent will obtain a college degree by Spring2015. Put another way, this optimistic scenario suggests that 27 percent of the 77,501 students in the full cohort will obtain a college degree by Spring 2015.[39]

One-year college persistence. Rather than college enrollment, we focus now on persistence to the end of the first year of college, as some students who enrolled as first-time students did not actually matriculate or dropped out before the end of one year.[40] Almost 80 percent of Fall 2009 college enrollees in this cohort remained in college through Spring, 2010—the end of their first year. Figure 5 indicates unadjusted one-year college persistence rates by student socio-demographic background. We find that roughly two out of three Asian and white high school graduates began college and persisted to the end of their first year, compared to less than half of black and Hispanic students. One out of three LEP and IEP students completed the first year of college, compared to roughly 57 percent of non-LEP and non-IEP students. Females

[38] See Cook, 2013.

[39] Two-thirds of students in the cohort graduated from high school, two-thirds of these high school graduates enrolled in college for Fall, 2009, and national estimates suggest that just under two-thirds of full-time students at four-year public institutions graduate within six years. Many of the students in this NYC cohort, however, will not have the luxury of attending full time, and many will also attend two-year institutions. National estimates suggest that fewer than one in three part-time students attending either public four-year or two-year institutions will graduate within six years (Cook, 2013).

[40] The one-year persistence rate presented here may differ somewhat from official DOE rates given our different analytic sample.

had a slight enrollment and persistence advantage over their male peers, as did free/reduced-price lunch students over non-FRPL students.[41]

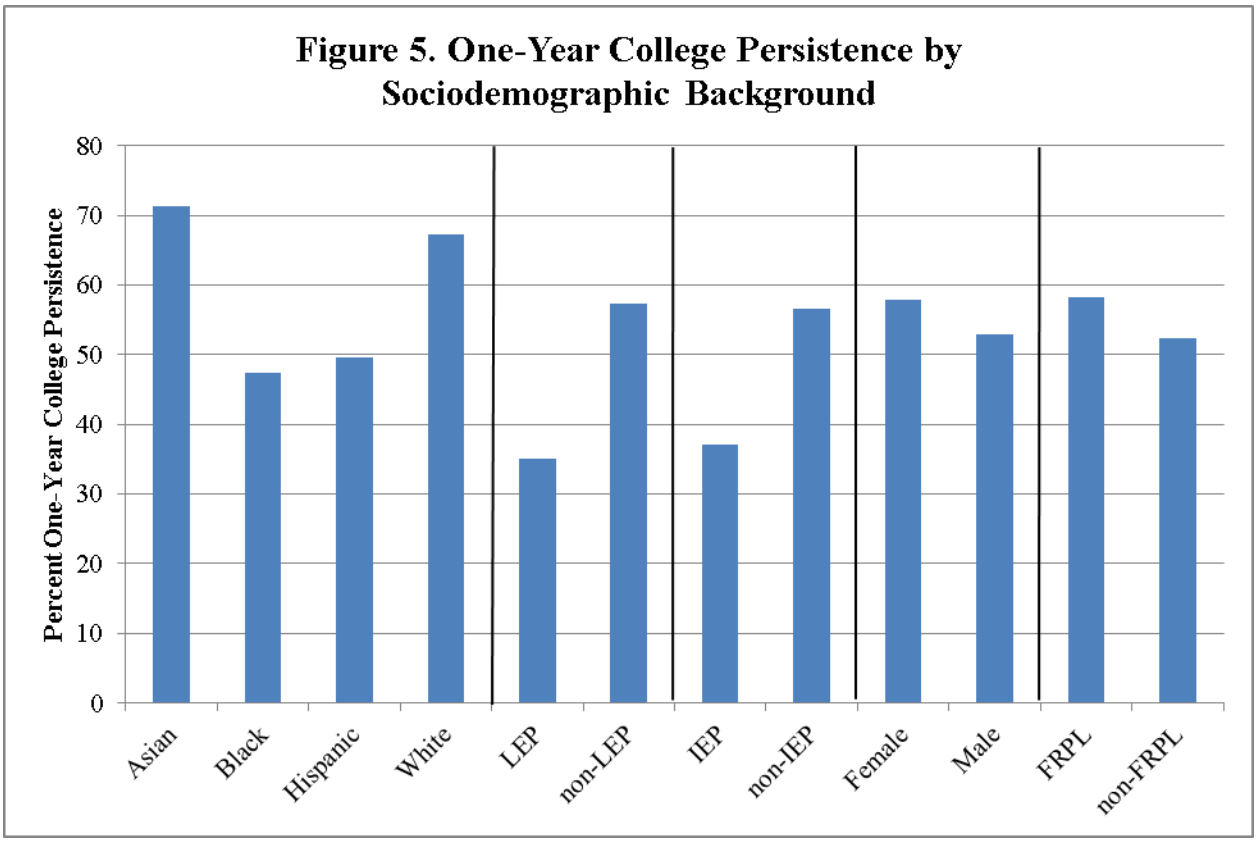
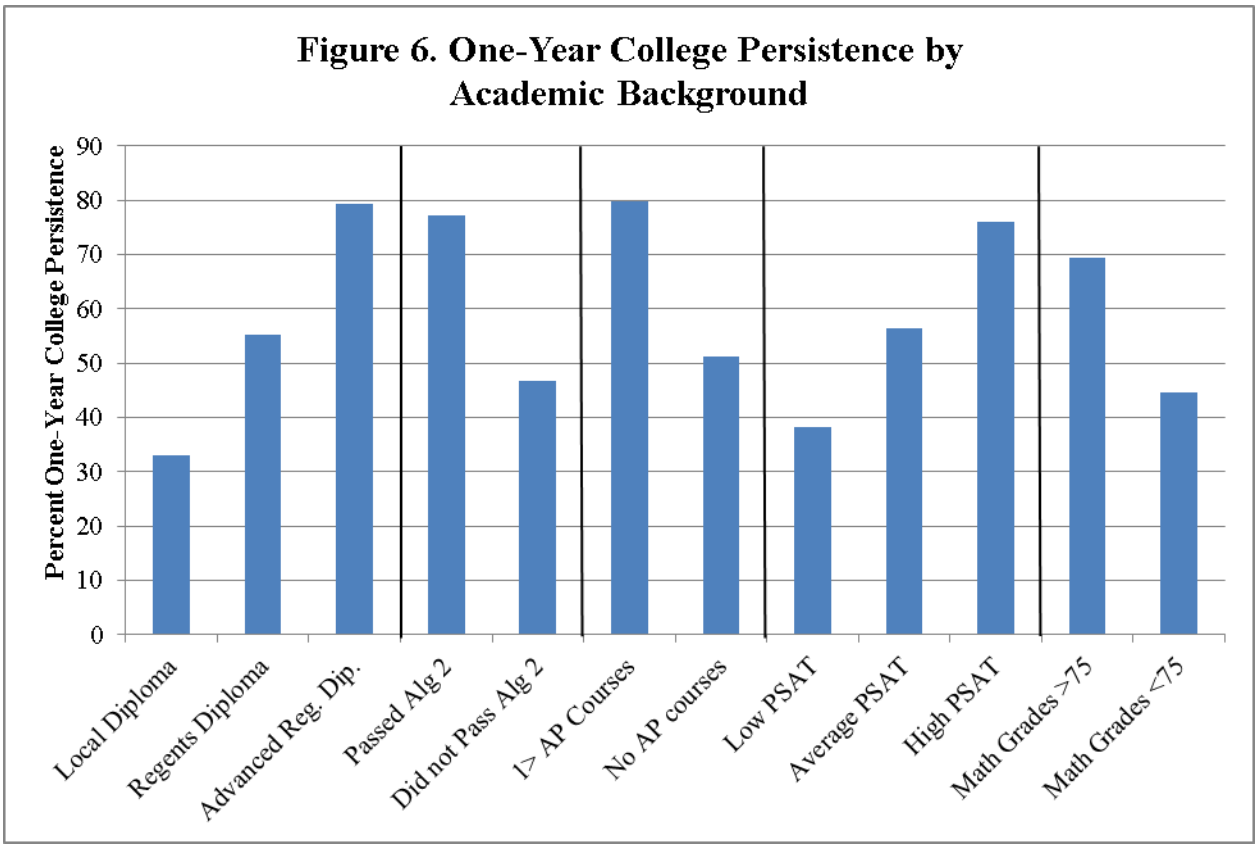


Figure 6 presents similar results organized by student academic background. Striking here are the differences in one-year college persistence by high school degree type. Fewer than one in three students who received a local diploma began college in Fall 2009 and remained enrolled at the end of that first year, compared to over 50 percent of Regents diploma holders, and almost 80 percent of those with an Advanced Regents diploma. More than three of four graduates who completed algebra 2 or beyond in high school remained in college after year one, compared to fewer than half of students who completed a less rigorous mathematics sequence. Similarly, 80 percent of students who passed at least one AP course persisted through one year of college, compared to roughly half of students who took no AP courses.[42] Unsurprisingly,

[41] Again, this suggests the weak nature of the FRPL measure.
[42] We should stress here the complicated associations between coursetaking and post-secondary outcomes. Students who complete advanced courses in high school typically have stronger academic

students with higher PSAT scores were more likely to enroll in college and complete the first year. Just over one in three students with low PSAT scores remained in college in Spring2010, compared to 57 percent of students with average PSAT scores and 76 percent of those with high PSAT scores.[43] In terms of high school grades, we find an enrollment and one-year persistence rate of almost 70 percent among students whose four-year math course grades were above 75 (roughly the average for graduates in this cohort), compared to less than 50 percent for students with below-average math grades.[44] Note that these proportions closely mirror those for passing/not passing Algebra 2.



abilities. The extent to which taking advanced courses actually improves academic performance is less clear.

[43] Scores are averages of the PSAT math, writing, and critical reading tests. Low scores are below -0.5 SDs, high scores are above 0.5 SDs, and average scores fall between -0.5 and 0.5 SDs.

[44] Our interest in grades here is driven somewhat by the fact that grades partly reflect student non-cognitive traits, such as persistence, which are also clearly important to college completion. But as we know from the above analyses regarding grade/Regents score differentials, grades do not reflect a constant and objective measure of student ability. However, math grades may be more consistent indicators of ability than grades in other subjects, given the more sequential nature of mathematics curricula.

Figure 7 then displays the adjusted associations between student characteristics and the likelihood of one-year college persistence. Even after accounting for a host of other student social and academic characteristics—including test scores and grades—we still find strong associations between high school diploma type and college persistence. Students who received an Advanced Regents diploma had odds of being enrolled in college after one year that were 60 percent greater than those of students who received a regular Regents diploma, while students who had obtained a Local diploma had odds of college enrollment and one-year persistence that were almost 50 percent lower than their Regents diploma holding peers. All else equal, both math grades and PSAT scores remained positively associated with persistence.[45] Asian students had somewhat greater odds of persistence compared to white students, while both black and Hispanic students were less likely to enroll and complete one year of college, even after controlling for key factors such as diploma type, test scores, and grades. Similarly, LEP students had adjusted odds of persistence that were roughly 33 percent below those of non-LEP students. IEP status was negatively (but weakly) related to persistence, while older students were somewhat less likely to persist.[46]

[45] A one-standard-deviation increase in math grades was associated with a 32.6% increase in the odds of college enrollment and one-year persistence, while a similar increase in combined PSAT scores was associated with a 17.8% odds increase.

[46] A one-year increase in age was associated with a 27.5% reduction in the odds of persistence.

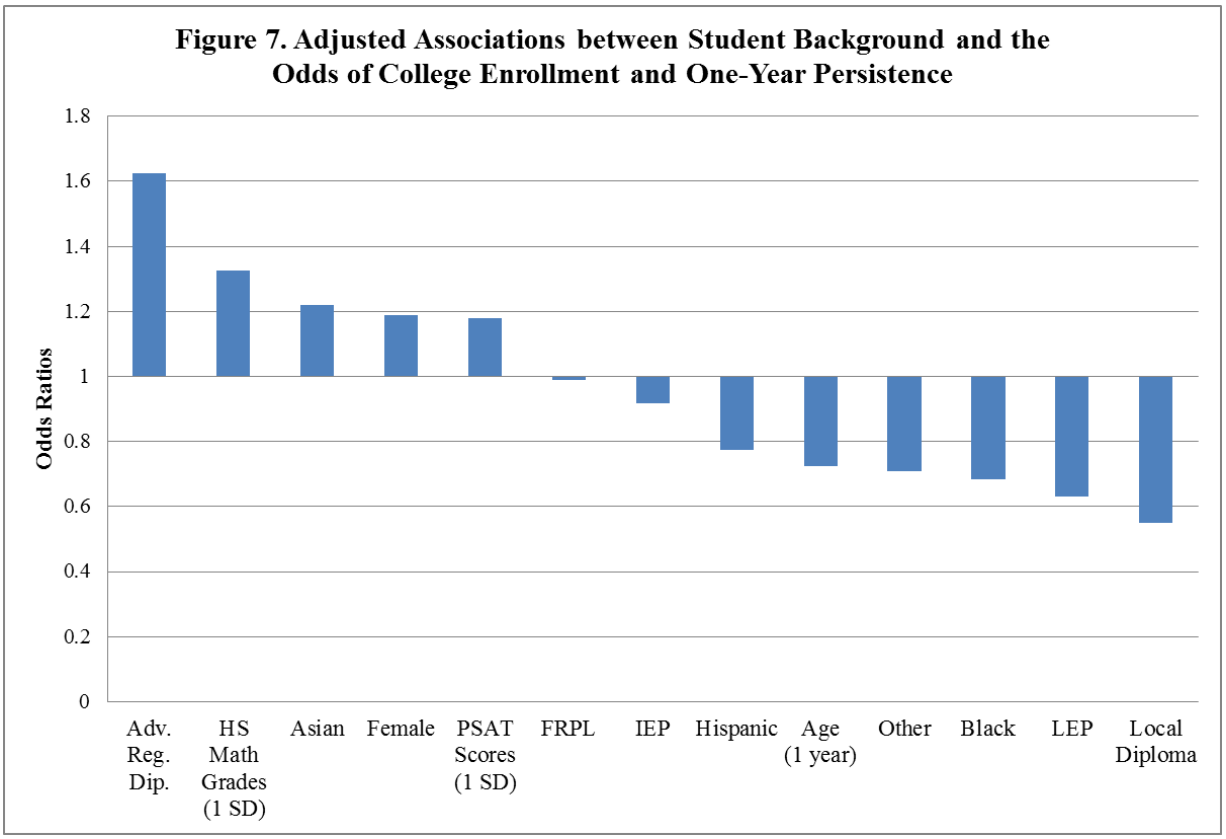


Figure 8 displays the results of four separate logistic regressions. The aim here is to gain a rough estimate of the extent to which four sets of student background measures explain differences in college enrollment and one-year persistence rates across students. Of these four measures (or groups of measures), whether a student received a local, Regents, or Advanced Regents diploma explained more variability in one-year college persistence than either PSAT scores, high school mathematics grades, or student social and academic background characteristics (including race/ethnicity, FRPL, IEP and LEP status, gender, and age). Note that student math grades tell us as much about the likelihood of one-year college persistence as do the full collection of student background characteristics combined.[47]

[47] We do not include both high school coursetaking and diploma type here, as they are in many cases mutually dependent. For example, 98% of students who received an Advanced Regents Diploma passed Algebra 2, and 86% of students who passed Algebra 2 received an Advanced Regents Diploma).

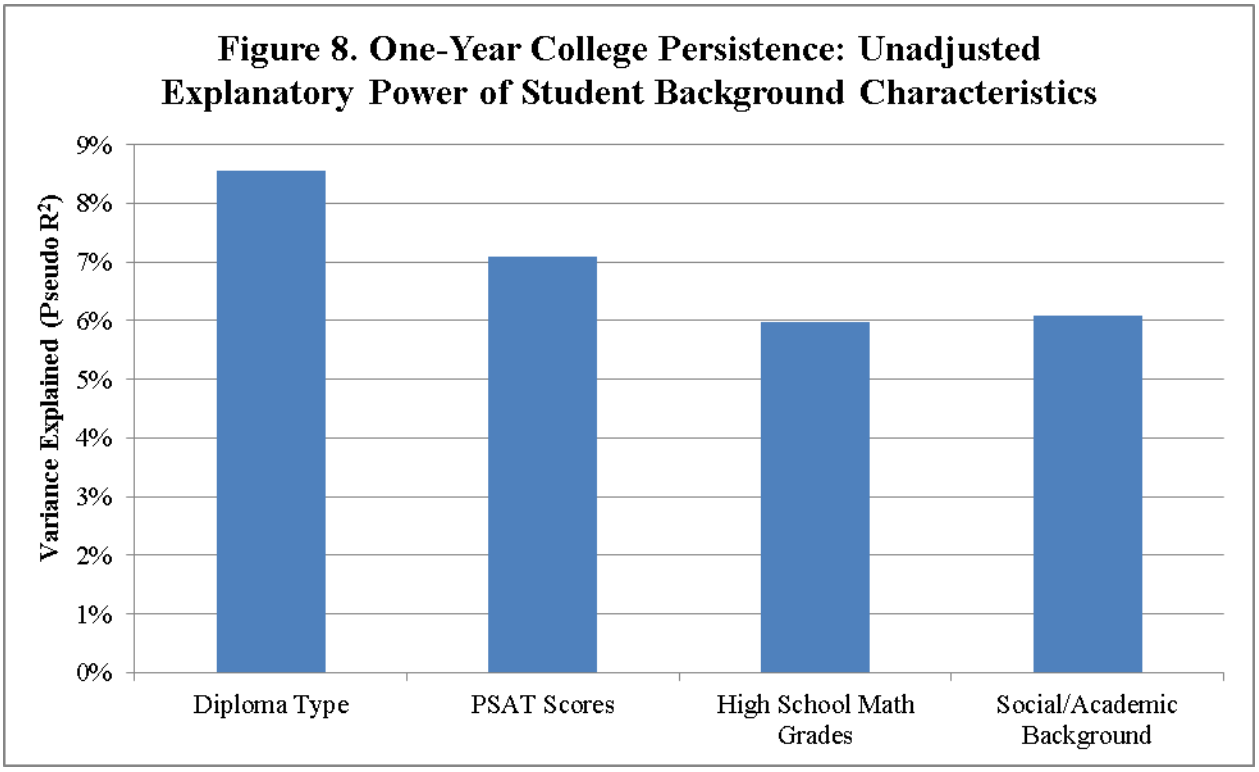


Figure 8. One-Year College Persistence: Unadjusted Explanatory Power of Student Background Characteristics
Variance Explained (Pseudo R²)
9%
8%
7%
6%
5%
4%
3%
2%
1%
0%
Diploma Type
PSAT Scores
High School Math Grades
Social/Academic Background

Conclusions and Implications

This report on the experiences of the class of 2009 suggests students in New York City are on different paths that lead to different ends. Some students—many more than in the past—are on a path to college. 67.1 percent of the students in this cohort who started high school in 2005 graduated from high school within four years, a substantial increase from a 51 percentgraduation rate for those who started high school in 1999.[48] Over two-thirds of these high school graduates enrolled in college for the 2009-2010 academic year, and almost 80 percentof those who started college in Fall 2009 remained through Spring 2010.

Despite promising developments like these, however, our findings also make clear that far too many students find themselves on pathways that do not lead to college or productive careers. These "dead end" tracks are reflected in the fact that students who failed to reach the third-grade ELA standard or accumulate at least 10 credits by the end of ninth grade had significantly lower odds of graduating in four years than their peers. Reflecting the stark differences in paths through the system, one in three students who attended higher-needs schools in elementary, middle and high school graduated within four years, while almost three of four students who attended no higher-needs schools graduated within the same period of time. In these "dead end" tracks, students were also less likely to experience the kinds of consistent opportunities to learn that they needed to reach graduation. In particular, black and Hispanic students and those with IEP's were more likely to change schools and were more likely to be absent and to miss school due to suspensions in both middle and high school. Furthermore, black, Hispanic, and IEP students were more likely to attend high schools with higher percentages of low-performing peers than other students, and these students typically completed fewer college-preparatory math courses.

While we caution again that these results reflect relational rather than causal associations, taken together with the results of similar studies in other cities like Chicago and with previous studies in New York, [49] these findings suggest that in many ways this one cohort of students experience two different school systems. For the group that is on or near the path to college, advanced classes, clear benchmarks and associated efforts to get and keep them "on-track" may serve as the scaffolding and support they need to make their way to high school graduation and

[48] See Kemple (2013).
[49] See for example, Kieffer, Marinell, & Stephenson (2011) and Allensworth & Easton (2007).

through the pipeline to college. While there are still too few black, Hispanic, LEP, and IEP students who are on the path to college, there are thousands of students from these groups who are on this path. For these students, this path to college can be seen as a system of supports, and drastic changes in the system as a whole could be seen as threatening and problematic. For those students who are already off the track for college when they begin school and make little or no progress during their school careers, however, benchmark assessments, advanced classes, Algebra 1, Regents exams, AP courses, AP exams and other graduation and college entrance requirements (and the disconnect between them) look more like a system of obstacles. Clearing one obstacle only leads to the next, and collectively they serve as the barriers that keep students out of the pipeline to college. For the students who struggle the most, dismantling the pipeline and making drastic changes in the system of obstacles may be required.

Implications for policy and practice

Given these drastically different odds and outcomes for students in New York City, the existing system needs to improve the support for those students who can see a path to success. Simultaneously, the system also needs reinvention to remove those predictable and systemic barriers that may not help or even hinder many students from making progress and getting anywhere close to a pathway to college. While many in education who have attempted to develop new schools and programs often describe their challenge as one of trying to build a plane while flying it, the dual demands described here suggest that the challenge may be even more complex: we have to learn how to fly the plane while we take it apart.

Building the pipeline to college. A number of education initiatives underway in New York City may already have contributed to the increases in high school graduation rates experienced by recent cohorts of public school students, and these initiatives may provide a foundation for further efforts to build and expand the pipeline to college in the future. For example, the transition to the Common Core Learning Standards (CCLS) has been a major focus of the New York City Department of Education. While that work is also in relatively early phases, it addresses some of the issues related to the inequitable access to rigorous coursework and the inconsistent expectations that these findings highlight. In particular, recent work in New York City has included pilot efforts to develop Common Core-aligned assessments and common

rubrics that could help to create greater consistency in learning experiences across classes and schools that could help to address the differentials between grades and Regents scores reported here.

Work on suspensions provides another example of initiatives that might be built upon. Although students from the class of 2009 who were suspended even once were much less likely to graduate than their peers, policies put in place after 2009 may be decreasing the number of suspensions experienced by subsequent cohorts of students. For example, in order to make the system more transparent and decrease the number and length of suspensions, in 2010 the NYC DOE and the New York City Council revised the discipline code and school safety policy by passing the Student Safety Act. In 2012 the NYC DOE took additional measures to limit the number and length of suspensions, including capping the number of days young children can be suspended for low to mid-level offences.[50] Future studies should track whether and how these and similar changes affect the graduation rates and other outcomes for subsequent cohorts of students.

In addition, some schools, including particular small schools and charter schools, have been in the vanguard of efforts to respond to the new demands of college and career readiness. The initiatives in these schools include a variety of efforts designed to focus explicitly on the issues of inequity, "on-track" performance, attendance and socio-emotional climate and support that are highlighted in this report. Many of these efforts, some of which are documented in the snapshots that are also a part of this project, are at a relatively early stage of development. Nonetheless, our findings do show some positive developments on some of the key measures that are related to high school graduation. Thus, students in the class of 2009 who attended these kinds of small, non-selective schools earned more credits in ninth-grade and missed fewer days of school in ninth grade compared to similar students who attended medium-sized and large non-selective high schools–schools that were the norm only a decade ago.

The development of many different kinds of schools in New York City also reflects one way of supporting multiple approaches to addressing the dramatically different needs of students who are already on the path to college as well as those who are struggling to get on it. Thus, charter schools and small schools like Uncommon Schools, Achievement First, and KIPP that

[50] For a recent discussion of the issue of suspensions and related initiatives see the report of the New York State Permanent Judicial Commission on Justice for Children (2013).

focus on college preparation might provide one set of approaches to getting students on-track; the work of schools, like the International High Schools, might represent an alternative that provides differentiated supports for students for whom the conventional pathway to college might look more like a system of obstacles; and the experiments of schools in the iZone might offer more radical approaches to rethinking the basic organization and supports for learning that students struggling the most might need. Given how soon many students fall "off-track," considering how these different kinds of approaches might help students throughout their school experiences should also be pursued. Supporting a variety of alternative approaches, however, requires a continuous, comprehensive effort to attend to issues like segregation and stratification and to ensure that there are no dead-end tracks and that all students have an equal opportunity to learn.

Dismantling the pipeline? Despite these promising developments, our findings suggest that for far too many students it is not sufficient to simply fix the conventional pipeline to college. For these students, conventional assumptions about school–when it should begin, how time should be used, how learning experiences are sequenced and structured—need to be re-examined.

Reconsider when the "system" begins. Findings like the already disturbing achievement gaps found in the earliest indicators of performance suggest a greater focus on early interventions. It is difficult to conceive of truly substantial improvements in the "whole system" if children do not enter the system until they are five or six years old. Initiatives like universal preschool that affect all students may be worth exploring. But some students, particularly black and Hispanic students, are low-performing early in their school careers and never reach college- and career readiness benchmarks. Initiatives focused on reaching those students at a young age may be particularly beneficial.

Rethink the focus on time and "credit." Most of the variability in the performance among New York City students is not between schools but within the same school. This suggests that how students are placed or end up in different courses, and how they accumulate credits once they are assigned to schools, are also critical areas for policymaking.[51] Given that the factors that inhibit credit accumulation include those like absences and suspensions that are related to

[51] While examinations of data at the level of individual teachers are beyond the scope of this report, assignment of students to teachers is another critical issue.

time in school, the kinds of extended learning time initiatives that are already underway may be particularly fruitful. At the same time, as many advocates for extended learning time point out, adding more time for conventional academic activities may not be particularly effective for students who have made little or no progress during the regular school day and are far behind in the number of credit hours they need to accumulate. If these students are out of school too often, to what extent can requiring them to spend more time in school address the problem? Thus, some rethinking of the underlying expectations around credit accumulation—themselves based on antiquated notions of seat-time—is in order. The necessity for students to follow particular sequences of units in a course and sequences of courses in a subject in a limited period of time also creates considerable challenges for students who are not in school continuously, particularly those like IEP students who may be changing schools frequently as well. As a consequence, the rethinking of the use of time in school needs to begin with the assumption that some students, particularly some of the students who are struggling the most, may be unlikely to attend school on a continuous basis.

Re-examine the basis for the organization of learning. This report highlights the problems that can come with inequitable and inconsistent expectations and reinforces the need to maintain clear, high standards for all. However, there is a complicated relationship between holding all students to high common standards and ensuring that students at all levels of development and performance experience appropriate support. For struggling students, repeatedly confronting demands for performance they cannot reach can undermine the motivation and confidence they must have to persist in school. Particularly as socio-emotional and other factors are recognized as critical indicators of readiness for success in college, the relationship between the rigor of academic expectations and support for all aspects of students' development should be examined. How can students who may be far below the academic expectations for their age or grade work toward those expectations at the same time that they participate in learning experiences that contribute to their personal development?

At least in part, this problem reflects the fact that schools continue to group students into conventional levels and grades that are based largely on age. The Common Core, for the most part, continues to reflect this organizational structure. Furthermore, students from different backgrounds often have a difficult time keeping pace within those structures. This can lead them to be sorted into lower level and specialized classes, which, in turn, can result in inequitable

learning opportunities. Rather than basing the organization of students and courses on students' ages, alternative groupings and structures for learning should be created based on the pace of students' development in cognitive and non-cognitive domains. Developing alternative structures, however, will also require re-examining the benchmarks, rewards, and other incentives that define "progress" in the current system. As expectations for college-readiness continue to expand to include less conventional cognitive and non-cognitive goals, policymakers will need to establish new indices and new accountability schemes that respond to the variety of students' developmental needs.

Key issues for future research

The findings presented here suggest that even as policymakers and practitioners in New York City work to respond to the new demands for college and career readiness, there are several key areas where additional research can help inform future policymaking. First, the experiences of subsequent student cohorts should be researched to corroborate and extend the results reported here. That research should strive to uncover continuing patterns of inequity as well as to examine whether the promising initiatives pursued so far are having the intended positive effects on all students.

The research described in this report provides a retrospective look at an entire cohort of students. A second area for research would be regular prospective documentation of the experiences of new cohorts as they enter and move through the system, which could help to flag early patterns of inequity and to identify key factors that may be contributing to those patterns. Although many data systems now focus primarily on in-school factors, the sophisticated geographic mapping tools now available make it possible to explore key relationships between the neighborhoods where students live, school factors, and student's experiences and outcomes. These community-school-student examinations provide a much better understanding of both the factors that help students to get on track and the factors that help explain why large numbers of students never get on track and show little or no progress in school. Are there some neighborhoods that have disproportionately high numbers of students who never experience much success in school and are there particular characteristics of neighborhoods that might explain those outcomes? Are kindergarteners from similar income levels but from more diverse neighborhoods more likely to experience success in their schools than their peers in more

segregated neighborhoods? Are students who stay in the same neighborhood during much of their school careers more or less likely to end up in low-performing schools, and are those students more or less likely to be ready for college or a career?

Third, in conjunction with ongoing documentation of the experiences of future cohorts of students and the examination of connections between community and school factors, research should look much more carefully at who students are. Our analyses suggest that the conventional broad categories of racial background may obscure important interactions that underlie many of the more general findings we report. For example, broad generalizations about students designated as LEP are likely to be unhelpful. As our analyses indicate, there may be different groups of LEP students who are having very different experiences in the New York City public schools. While LEP status was associated with slightly fewer absences and higher credit accumulation than their peers, Hispanic students—who include a high percentage of LEP students—had more absences and lower credit accumulation. In contrast, Asian students—another group with high percentages of LEP students—had lower rates of absences and higher credit accumulation compared to the district average. To improve practice and policy, LEP status may need to be disaggregated into more precise descriptors, taking into account additional characteristics. These might include race/ethnicity, immigrant status (first or second generation), or, for first generation immigrant students, number of years of schooling before coming to the United States. Geographic mapping could be used to look at characteristics such as where students live and how often they move once they come to New York City. Further analyses could then explore interactions more deeply to pinpoint the successes and challenges that different students experience. Tracking how different groups of LEP and immigrant students move through different neighborhoods and through the school system—and the extent to which they are sorted into different schools and classrooms—may provide a crucial window on their learning experiences.

General categories, like students with IEP's, ignore critical variations that could help to explain the kinds of initiatives and supports that might be most useful for students within this population. Regularly documenting the performance of student cohorts could be accompanied with survey and interview-based studies, perhaps engaging the students themselves, in studying what helped students from different backgrounds as they made their way through the system.

Fourth, in addition to developing more sophisticated methods of exploring the connections between students, where they live, and their paths to college, this report suggests several key long-term issues that deserve particular attention. Given the central place of school choice in New York City, an effort should be made to provide options that do not cause further segregation and academic stratification. While specialization into particular tracks and schools can provide more differentiated support for students, that same differentiation and specialization can then make it difficult for students who have started down one path to move to another. However, recent research suggests that there are design decisions that can lead to approaches that support more and less mobility between and among tracks.[52] These include the *choice procedure* associated with the initial selection of schools and courses; the *transparency* of the procedure and its consequences; and the *permeability* of the system—the extent to which changes between tracks are allowed and encouraged. Thus, systems like those in New York City in which students are often tracked (implicitly and explicitly) relatively early and where the differences between tracks and outcomes are often unclear to some parents and students can contribute to more rigid tracking and less opportunity for mobility between tracks which can exacerbate inequities. In contrast, in systems like Finland, choices among specialized schools and tracks can be made relatively late in students' school careers, the options are relatively clear, and students are given opportunities to change schools and tracks and to "catch-up" through additional coursework and other means should they need to.

The analyses presented here suggest that research should explore different possibilities for choice procedures in middle and high school; ways to make the choice procedure and the consequences of those choices clearer; and experiments with waivers or other means that allow students who might not normally make it into advanced or specialized tracks to do so. Along the same lines, research should explore efforts to hold students to high common standards and support all aspects of their development at the same time. Are there some requirements or benchmarks for accumulating credits, for graduating or for getting into college that can be removed without decreasing students' chances of success? For example, in order to be successful in college, do recently arrived students who speak little or no English but are doing well in rigorous coursework and persisting in high school need to able to pass Regents exams that require high levels of understanding of English in order to succeed in college?

[52] See Von reis Saari (2013).

Final Considerations

The questions raised here reinforce the need for policies and practices that provide clear, transparent paths that reflect students' different developmental needs but that all lead to the same place: success in college and beyond. While we have outlined some specific steps that may help achieve this goal, the analyses of the experiences of the class of 2009 also suggest two general considerations to guide policymaking moving forward.

Policymaking should be systemic. By following a single cohort of students, this analysis highlights the extent to which the issues students face extend across individual classrooms and schools and require systemic reforms. However, systemic approaches to reform often focus on scaling-up solutions to all relevant units in the system at a particular moment in time. In contrast, these findings provide a reminder that students do not experience the system all at once; students experience the system over thirteen years or more as they and subsequent generations develop and grow. That suggests that a generation of interventions might be needed to address systemic issues like the extent of stratification between and within the schools. Nonetheless, the system will always encompass students who are just starting their school careers and those who are nearing the end and cannot wait for a generation of initiatives to take effect. As a consequence, systemic policy-making needs to balance both the long-term and short-term needs of current and future generations of students as well as the needs of students who are well-served by the system and those who aren't. Ultimately, individual success stories of students, teachers, programs and schools need to be documented, and we need to learn from those experiences. But it is the rare occasion when these local solutions apply across all students and contexts. More than focusing on what it takes for small numbers of schools and students to beat the odds, we need a systemic approach to policy and practice that changes the odds so success is the expectation for all rather than an exception to the norm.

Policymaking should be collaborative. This report makes clear the competing interests that are at work in education today. Correspondingly, the systemic nature of the problems and the persistence of the inequities that particular groups of students experience mean that individual and individualized approaches cannot be pursued in isolation. Different approaches are always part of the collective work in and on the system—work that has a significant impact on the segregation and stratification of students, the distribution of resources, and the equity of

outcomes for all. While the findings and implications explored here suggest that there may be no "one best system," we hope that this report illuminates some of the possibilities for a better system that responds to the rights and needs of every student.

References

ACT (2006). *Ready for college and ready for work: Same or different?* Iowa City, IA.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.

Allensworth, E., & Easton, J.Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Chicago: Consortium on Chicago School Research.

Allison, P.D. (2002). *Missing Data.* Thousand Oaks, CA: Sage University Paper No. 136.

Bloom, H., Thompson, S.L., & Unterman, R. (2010). *Transforming the High School Experience How New York City Small Schools Are Boosting Student Achievement and Graduation Rates.* New York: MDRC.

Bloom, H. & Unterman, R. (2012). *Sustained positive effects on graduation rates produced by New York City's small public high schools of choice.* New York: MDRC.

City University of New York (2013). *Performance management report 2012-13: Year-end University Report, preliminary*. New York: Author. Retrieved from: https://www.cuny.edu/about/administration/offices/ira/ir/data-book/current/accountability/PMPUniversityReport_2012-13_Prelim.pdf

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Kemple, J. (2013). *The condition of New York City high schools: Examining trends and looking toward the future*. New York: The Research Alliance for New York City Schools.

Kieffer, M.J., Marinell, W.J., & Stephenson, N.S. (2011). *The middle grades transition study: Navigating the middle grades and preparing students for high school graduation*. New York: The Research Alliance for New York City Schools.

Lee, V.E., & Ready, D.D. (2009). U.S. high school curriculum: Three phases of contemporary research and reform. *The Future of Children, 19*(1), 135-156.

Nathanson, L., Cole, R., Kemple, J., Lent, J., McCormick, M., & Segeritz, M. (2013). New York City School Survey 2008-2010: Assessing the reliability and validity of a progress report measure. New York: The Research Alliance for New York City Schools.

Nathanson, L. Corcoran, S., & Baker-Smith, C. (2013). *High school choice in New York City*. New York: The Research Alliance for New York City Schools.

New York Civil Liberties Union (NYCLU) (2011). *Education interrupted: The growing use of suspension in New York City's public schools*. New York: Author.

New York City Department of Education (2011). Final changes to the 2010-11 Progress Report. Retrieved from: http://schools.nyc.gov/NR/rdonlyres/2C66C50A-4CEA-4DEB-845A-F1C04474754A/0/Final_Changes_Presentation_2011_11_04.pdf

New York City Independent Budget Office (2011). *New York City public school indicators: Demographics, resources, outcomes*. New York: Author.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in bservational studies for causal effects. *Biometrika, 70*, 41-55.

Rubin, D.B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127*, 757-763.

Reardon, S.F., Yun, J.T., & Eitle, T.M. (2000). The changing structure of school segregation: Measurement and evidence of multiracial metropolitan-area school segregation, 1989-1995. *Demography, 37*(3), 351-364.

Reardon, S.F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology, 32*(1), 33-67.

Rubin, D.B., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics, 52*, 249-264.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton Mifflin.

Von Reis Saari, J. (2013). *Moving on up? A framework for evaluating equality of access in education, with illustrations from Finland, Sweden and the United States*. Paper presented at the Annual Conference of the American Educational Research Association.

Appendix A

Data

These analyses employed longitudinal student-level administrative data supplied by the New York City Department of Education, supplemented with school-level data from the New York State Department of Education, the National Center for Education Statistics' (NCES) Common Core of Data (CCD), and the New York City High School Directories. We drew analytic samples from a base sample of 77,501 students the NYC DOE classified as being in the 2009 graduating cohort (based on expected rather than actual graduation year). The exact sample sizes for the K-8 analyses varied depending on the analyses, as students not enrolled in a New York City school in a given year were not used in the K-8 analyses, and were not treated as having missing data. The analyses of ninth-grade outcomes used a sample of 70,130 students who were in a New York City high school in ninth grade and were not attending a school associated explicitly with the criminal justice system, drug/alcohol rehabilitation, teen pregnancy, or mental health. The larger number of students lost, however, stems from the fact that not all of the 77,501 students in the base sample were enrolled in a New York City high school in ninth grade. The DOE organized these students into this cohort once they were admitted in 10th, 11th, and in a few instances, 12th grade, based again on their expected graduation year.

The sample used to explore student mathematics coursetaking excludes students who transferred schools during high school, whether they left the New York City public school system entirely, or simply transferred among schools within the system. Transfer students pose a problem because their accomplishments are not attributable to one school alone. This is particularly a concern for our outcome of interest, progress through the math curriculum, because it is measured after the completion of a full four years of high school. Students who drop out of high school also pose problems for analyses of school effects because they do not receive the full four-year "treatment" of attending high school. Thus, students who drop out during their first four years of high school are also excluded.

However, excluding transfer students and dropouts is clearly problematic as well, because it eliminates the most disadvantaged and academically underprepared students. Moreover, because students who dropped out tended to fall lower on the achievement

distribution, excluding them may unfairly penalize schools with higher retention rates and reward schools with higher dropout rates. Thus, in Appendix B Table 7, we re-analyze our models of students' progression through the math curriculum on a sample that includes students who transferred within New York City public schools and students who dropped out. Students who transferred out of the New York City public school system are still excluded because we do not have access to information on their most advanced math course taken in four years of high school. For these models students are nested in their first high school. These models assess the most advanced level of math that students reach regardless of whether they dropped out. Thus, a student who passed Algebra her freshman year and then dropped out would be given credit for reaching Algebra. Put simply, this technique measures students' exposure to math during high school, regardless of how long they remained in school. We then controlled for indicators of whether students dropped out and whether they attended more than one high school. Because these results largely reflect the same relationships as those presented in Table 6, we refer to them only to reference those elements that do differ.

Analytic Methods

The ninth-grade analyses of absences and credits earned used hierarchical linear modeling (HLM) which analytically nested students within schools (see Bryk & Raudenbush, 2002). The analyses predicting high school graduation were conducted using a logistic regression that did not nest students in schools, given that so many students changed schools during high school. The fact that this model employed a single dichotomous indicator or graduation precluded us from constructing cross-classified models, which would have allowed us to nest students within multiple schools over time. We also considered hazard/survival models for the graduation analyses, but our understanding of the literature on high school dropouts suggests that leaving school is far from a discrete event, and that it does not occur at a specific point in time. The models of high school mathematics coursetaking used the multi-level modeling capabilities within STATA 12 (xtmixed). We provide further information below about several school-level measures that likely require further information. For more information (or suggestions) regarding model specifications, please contact the first author.

School academic selectivity. In the high school analyses, all schools are categorized as either academically selective or non-selective. For single program schools the designation is

clear-cut—the school is categorized based on the selectivity of its program. For multi-program schools, categorization is more complex. For these schools, we follow the approximate method of classification utilized by Quint et al. (2010). For each year the number of seats available in schools' academically selective programs are calculated as a proportion of the total number of enrolled first-time ninth-grade students. If the proportion of all first-time ninth graders enrolled in academically selective programs was greater than or equal to 50 percent, the school as a whole was considered to be academically selective for that year. Information on program selectivity and number of seats was taken from the New York City High School Directory from the subsequent school year.

School surveys. Several sets of analyses also use four school-level factors associated with the NYC DOE's school survey, which is completed annually by students, parents, and teachers. These measures touch on four broad areas of school climate, including safety and respect, communications, engagement, and academic expectations. Because the surveys were not yet conducted when this student cohort was in ninth grade (2005-2006), we instead use the survey reports from the 2006-2007 academic school year. The analyses of high school mathematics coursetaking take the average score for each factor across the 2006-2007, 2007-2008 and 2009-2009 school years. Within our school sample, the correlations within factors over years ranged from .80-.85 for safety and respect, to .66-.74 for engagement, suggesting that these measures were relatively consistent over time within schools. For a full and helpful discussion regarding the content, reliability and validity of items on the New York City school surveys, and the resulting factors, see Nathanson, Cole, Kemple, Lent, McCormick & Segeritz (2013).

Propensity scores. To obtain the estimate of the effect of suspension on the likelihood of graduation, we employed propensity score methods. Multi-level propensity score matching allows us to create a more plausible counterfactual for suspended students than all students who were never suspended. With this technique, we use student characteristics to calculate the predicted probability of being assigned a suspension, and then match non-suspended students with suspended students based on similar probabilities (Gelman & Hill, 2007; Rubin, 1997; Rubin & Thomas, 1996). In this technique, the predicted probabilities for suspension assignment are calculated using logit models, with the predictors being the confounding covariates that are pre-treatment characteristics related both to the likelihood of suspension and to the long-term

outcome in question. To account for the nested nature of students within schools, we forced an exact match on students' ninth grade school.

Once calculated, we use propensity scores as weights in a logistic regression model to estimate the average effect of the treatment on the treated. In other words, we estimate how suspended students would have performed, on average, had they not been suspended. This technique allows us to provide more valid estimates over traditional regression techniques, as it ensures that the treatment and control group are balanced and have substantial overlap (Gellman & Hill, 2007; Rosenbaum & Rubin, 1983). Causal interpretation of these results should still be used with hesitation, however, as we are only able to account for observed confounding covariates when calculating the propensity of being assigned a suspension (Shadish, Cook, & Campbell, 2002).

Missing data. As with most large administrative datasets, students in our samples were missing data on particular measures, and it is likely that these data are not missing completely at random. By far the largest concern was with missing eighth-grade test score data, largely for students who did not attend middle school in New York City. All multivariate models employ multiple imputation to address missing data, using the MI package in STATA to generate five imputed datasets. In general, multiple imputation holds several advantages over other common strategies for addressing missing data. This approach allows for use of the full dataset, conserving all valuable information, and also produces less biased estimates (Allison, 2002; Rubin, 1996). We estimate our results separately on each of five imputed datasets, and then combine the results according to Rubin's rule (Rubin, 1987). In our imputation models we include all variables that are used in later analyses.

Appendix B

When comparing the results presented in Table 6 to those presented here in Table 7, which re-estimate the same models on a broader student sample, including both within-district transfers and dropouts, the significance and general magnitude of the coefficients remains largely consistent. In general, it appears that the school-level effects are slightly reduced in the broader sample. Moreover, the indicators for dropping out and attending more than one high school are both strongly negatively related to students' progress through the math pipeline. Unsurprisingly, students who dropped out progresses more than a year and a half less far through the math curriculum than their counterparts who remain in school, on average. Similarly, students who attend more than one New York City public high school progress roughly one-half year less far than their peers who persist in the same school all four years, all else equal.

Table 7. Multi-Level Models of Highest Math Course Completed[a] (*n*=57,343)

	Model 1	Model 2	Model 3	Model 4[b]
Socio-Demographic Characteristics				
Asian[c]	0.52***	0.29***	0.27***	0.27***
Black	-0.44***	-0.22***	-0.17***	-0.17***
Hispanic	-0.43***	-0.27***	-0.23***	-0.23***
Native/Multi-Ethnic	-0.22***	-0.15*	-0.12*	-0.12*
Female	0.20***	0.02***	0.15***	0.15***
Free/Reduced-Price Lunch	0.01	-0.05***	-0.06***	-0.06***
Special Education	-0.84***	0.06**	0.031	0.03
English Language Learner	-0.12***	0.01	0.021	0.02
Behavioral/Academic Characteristics				
Overage for 8th grade		-0.03*	-0.06***	-0.06***
Middle school mobility		-0.05***	-0.05***	-0.05***
Entering math achievement[d]		0.48***	0.43***	0.43***
Entering ELA achievement[d]		0.15***	0.10***	0.10***
Days absent in 8th grade		-0.02***	-0.02***	-0.02***
Days tardy in 8th grade		-0.01***	-0.01***	-0.01***
Dropout		-1.63***	-1.65***	-1.65***
Attended >1 NYC DOE high school		-0.51***	-0.50***	-0.50***
First Math Course[e]			0.55***	0.54***
School Characteristics				
School Type[f]				
Small, Selective				-0.09
Midsize, Selective				-0.16*
Midsize & Large, Nonselective				-0.43***
Large, Selective				-0.27**
School offers Calculus				0.17**
% credited courses below Algebra				-0.01*
Within-school SD most adv course[d]				0.09**
Academic expectations[d]				0.24**
School safety[d]				-0.06
Communications[d]				-0.06
Student engagement[d]				-0.02
Intercept	3.55***	3.60***	3.60***	3.65***

*p<.05, **p<.01, ***p<.001

[a] The most advanced math course passed in four years of high school (0=No math, through 8=Advanced topics beyond Calculus)

[b] The following non-significant school-level predictors were removed in the interest of model parsimony: borough indicators, the percent of courses classified as prep/support, whether the school offers AP/IB math, school communication, student engagement, the percent of students that were FRPL eligible, special education, LEP, or overage.

[c] All racial/ethnic groups are compared to white students.

[d] Measure is z-scored (M=0; SD=1)

[e] Level of first math course taken in high school (0=No math, through 8=Advanced topics beyond Calculus)

[f] School type categories are compared to small, nonselective schools.

PHILANTHROPY NEW YORK

Education Funders Research Initiative (EdFunders) is a project of Philanthropy New York and foundations associated with its Education Working Group. The foundations associated with EdFunders have diverse missions and vastly different strategies for implementing education reform, but share the goal of improving education in New York City using fact-based decision making.

EdFunders has invested in thoughtful, nonpartisan, accessible research with the aim of identifying and implementing education strategies that improve college- and career-readiness.

Contact:
Michael Hamill Remaley
Philanthropy New York
212.714.0699, ext. 222
mremaley@philanthropynewyork.org