Small High Schools and Big Inequalities: Course-taking and Curricular Rigor in New York City

Miya T. Warner

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013
This study examines whether small high school reform in New York City has fulfilled its goal of providing disadvantaged students access to rigorous mathematics curricula, thereby increasing their college readiness. Between 2002 and 2010 in New York City, 27 large, comprehensive high schools were closed or downsized and replaced by over 200 new small schools (Jennings & Pallas, 2010). Although extant research indicates that these schools have produced higher attendance and graduation rates (Bloom et al., 2010; 2012), the literature on small high school reform and college readiness remains inconclusive. To address this gap in the literature, my dissertation employs a longitudinal database of New York City student and school-level data from 2000-2010 to examine the impact of small high school reform on student math course-taking for two cohorts of students (the class of 2009 and 2010). I address the threat of selection bias by utilizing several propensity score matching techniques within a multilevel modeling framework.

I find a small, positive impact of attending a new, small high school on students’ progress through the math curriculum (one-sixth of a year) for the class of 2009, but not for the class of 2010. Yet while students in the new, small high schools, who are among the most disadvantaged in the city, might be faring slightly better than they would have had they attended an alternate high school option, they are still failing to complete even one semester of Algebra II/Trigonometry—the lowest level of course deemed “college preparatory” by the district. Furthermore, small high schools are not equally beneficial for all types of students. Black and Hispanic students appear to do better in the small schools than in alternate high school options, while the reverse is true for whites. Meanwhile, students with initially low math achievement benefit from attending small high schools, while students with middle-to-high levels of initial math achievement are better served elsewhere.
Moreover, the new, small high schools are much less likely to offer advanced math courses such as calculus or any Advanced Placement or International Baccalaureate math—effectively cutting their students off from the opportunity to take these courses. Finally, my results suggest that the rigor of math courses in the new, small schools may be weaker than in the alternate high school options in New York City.

Taken together with the existing research, my results suggest that the consequences of small high school reform in New York City are both more complicated and less positive than the reformers promised or district officials will admit (Gates, 2005; Walcott, 2012). While these schools are unquestionably improvements over the large, failing schools they replaced, they remain at the bottom of an intensely academically stratified school system, and they have failed to raise students’ college readiness in math. Moreover, these schools are particularly under serving high achieving students by cutting them off from rigorous, advanced math courses.
# TABLE OF CONTENTS

**CHAPTER ONE:** A POPULAR REFORM AND AN UNTESTED HYPOTHESIS ..................1

**CHAPTER TWO:** REVIEW OF RELEVANT LITERATURE ...........................................23

**CHAPTER THREE:** CONCEPTUAL FRAMEWORK ......................................................62

**CHAPTER FOUR:** METHODOLOGY .................................................................88

**CHAPTER FIVE:** SMALL SCHOOLS IN A STRATIFIED SYSTEM .........................117

**CHAPTER SIX:** “CONSTRAINED DOWNWARDS”: NEW SMALL SCHOOLS FOSTER EQUITY BUT NOT EXCELLENCE .................................................................145

**CHAPTER SEVEN:** CONCLUSION AND DISCUSSION ........................................176

**REFERENCES** ................................................................................................193

**APPENDIX A:** ADDITIONAL TABLES AND FIGURES ......................................209
LIST OF TABLES

Table 4.1: Curriculum Equivalents of the Math Pipeline Measure

Table 5.1: School Location and Creation Era, Differentiated by Size, Selectivity, and Freshman Cohort

Table 5.2: Socio-Demographic and Achievement-Based Characteristics of First-Time Ninth-Grade Students Enrolled in DOE High Schools by Size and Selectivity, 2005-2006 School Year

Table 5.3: Academic Composition and Curricular Structure of DOE High Schools Serving the 2005-2006 Freshman Cohort by School Size and Selectivity

Table 5.4: Group Means of Student Characteristics, Differentiated by the Highest Math Course Completed

Table 5.5: Structural and Compositional Characteristics of High Schools, Differentiated by the School Average of Students’ Highest Math Course Completed

Table 6.1: Multilevel Psychometric Properties of the Math Pipeline Measure

Table 6.2: Student-level Multilevel Models of Highest Math Course Completed, 2005-06 & 2006-07 Cohorts

Table 6.3: Between-school, Multilevel Models Exploring the Relationships Between School Factors and Students’ Progress Through the Mathematics Curricular Pipeline, 2005-06 and 2006-07 cohorts

Table 6.4: Between-school, Multilevel Models Exploring the Relationship Between Attending a New, Small, Nonselective School and Progress Through the Mathematics Curricular Pipeline, 2005-06 and 2006-07 cohorts

Table 6.5: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2005-06 cohort

Table 6.6: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort
Table 6.7: Multilevel Psychometric Properties of Math A and Math B Regents Scores and Performance Differentials…………………………………………………………………………………………165

Table 6.8: Average Course Marks, Regents Marks, and Performance Differentials of DOE High Schools Serving the 2005-2006 and 2006-07 Freshman Cohorts by School Size and Selectivity….168

Table 6.9: Between-school, Multilevel Models Exploring the Relationships Between School Factors and Students’ Performance Differentials, 2005-06 and 2006-07 cohorts………………………………………...172

Table 6.10: Between-school, Multilevel Models Exploring the Relationship Between Attending a New, Small, Nonselective School and Students’ Performance Differentials, 2005-06 and 2006-07 cohorts…………………………………………………………………………………………...174

Table A.7: Socio-Demographic and Achievement-Based Characteristics of First-Time Ninth-Grade Students Enrolled in DOE High Schools by Size and Selectivity, 2006-2007 School Year …….214

Table A.8: Academic Composition and Curricular Structure of DOE High Schools Serving the 2006-2007 Freshman Cohort by School Size and Selectivity……………………………………...215

Table A.9: Group Means of Student Characteristics, Differentiated by the Highest Math Course Completed, 2006-07 cohort………………………………………………………………………...216

Table A.10: Structural and Compositional Characteristics of High Schools, Differentiated by the School-average of Students’ Highest Math Course Completed, 2006-07 cohort…………………217

Table A.11: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort; Newly opened schools excluded……………………………………………………………………...218

Table A.12: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2005-06 cohort; Drop-outs excluded……………………………………………………………………...218
Table A.13: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort; Drop-outs excluded…........................................................................................................218

LIST OF FIGURES

Figure 3.1. Conceptual model of how school size influences student course-taking and educational outcomes in math.................................................................80

Figure 6.1: The relationship between race/ethnicity, attending a NSN school, and math course-taking..................................................................................................................159

Figure 6.2: The relationship between entering math achievement, attending a NSN school, and math course-taking ........................................................................................................159

Figure A.1: Highest mathematics course completed in four years (2005-06 & 2006-07 cohorts; includes drop-outs) ..................................................................................209

Figure A.2: Highest mathematics course completed in four years (2005-06 & 2006-07 cohorts; drop-outs excluded) ..................................................................................209

Figure A.3: Propensity score balance diagnostics (2005-06 cohort; multilevel PSM & PSW)…..210

Figure A.4: Propensity score balance diagnostics (2006-07 cohort; multilevel PSM & PSW)…..211

Figure A.5: Propensity score balance diagnostics (2005-06 cohort; exact match on middle schools) ..................................................................................................................212

Figure A.6: Propensity score balance diagnostics (2006-07 cohort; exact match on middle schools) ..................................................................................................................213
ACKNOWLEDGEMENTS

Writing this dissertation has been one of the most challenging and rewarding undertakings of my life. I could not have done it without the amazing constellation of family, friends, and mentors who have been in my corner from day one, and whose love, support, and advice have sustained and inspired me throughout my doctoral studies.

I have had the great fortune of working with two brilliant faculty advisors: Amy Stuart Wells and Doug Ready. These two scholars are very different from one another in many ways, ranging from methodology (qualitative v. quantitative) to their view of deadlines (squishy v. firm), but the attributes they share are even more important: a dedication to rigorous research and a deep resolve to remedy educational inequalities. Amy, you have pushed me to think deeper and more theoretically than I ever have in the past. Doug, if it wasn’t for you, I would still be scared of statistics and certainly would never have undertaken a quantitative dissertation! Thank you both for investing so much time in my development as a scholar.

I would also like to extend a special thanks to Aaron Pallas, who has given me invaluable feedback and advice throughout my studies, and who has dedicated more time and thought to my dissertation than I could have ever hoped for. I am also indebted to the rest of my dissertation committee: Luis Huerta and Carla Shedd for their time and support of my research.

A hearty thanks goes out to Elizabeth Chu and Cassandra Nader for their help with the interminable job of coding math courses. Really—thank you! I would also like to acknowledge Teachers College and MDRC for providing me with fellowships that helped support this work.

I could not have reached this point without my network of colleagues, whose critical feedback, support, and solidarity made weathering the highs and lows of the doctoral process much easier. A big thank you to the sociology team: Allison Roda, Terrenda White, Jacquelyn Duran, Bianca Baldridge, Lauren Fox, Kathy Hill, Hester Earle, and Richard Lofton. I would also like to
thank the talented and inspiring women of the DWG: Brooks Bowden, Mina Dadgar, Michelle Hodara, Alex Holod, Rachel Rosen, and Megan Silander. I will continue to count on you all for advice long after we have all graduated.

I owe a heartfelt thank you to my family and friends, for their unending and incalculable love and support. You know who you are. Without your love, generosity, encouragement, …and nagging, I would never have made it this far.

And finally, I would like to thank Cleveland. From unconditional love, to tough love, and from teasing to bribes, you have employed everything in your arsenal to help me finish. As this chapter closes, I know the most rewarding ones are yet to come. One team, one dream!
DEDICATION

To my first and most influential teachers—my parents. Every day I grow more grateful to you both for filling my childhood with love, curiosity, happiness, and a bit of magic. Your contributions to the world, your love for your family and one another, and your unending enthusiasm for learning continue to inspire me.
Chapter 1

A Popular Reform and an Untested Hypothesis

Introduction

Although economists predict that workers without college degrees will be at an even larger disadvantage in the 21st Century labor market (Carnevale, Smith, & Strohl, 2010), too many high school students are graduating unprepared for the rigors of college-level coursework. Even among those students who matriculate to higher education, roughly one-third need remediation in a core subject. Furthermore, those requiring even one remedial course in college have a substantially lower likelihood of graduating (Adelman, 2004; Bailey, 2009; Parsad & Lewis, 2003; McCabe, 2000). This disconnect between what is learned in high school and what is necessary for success in higher education is particularly pronounced for low-income and minority students, who take, on average, fewer advanced courses during high school than their more advantaged peers (ACT, 2006).

This is not a new problem. Getting students to complete more challenging academic courses has been a focal point for secondary school reform since the 1983 publication of *A Nation at Risk* disparaged the academic quality of the American high school (National Commission on Excellence in Education, 1983; Lee & Ready, 2007). Over the intervening decades, reform efforts have resulted in increased state graduation requirements, high-stakes graduation exams, a powerful standards and accountability movement codified in the federal No Child Left Behind Act, and the more recent Common Core Standards Initiative.

Another related policy strategy to boost college and career readiness, particularly among disadvantaged students in urban contexts, is small school reform, focused on reducing the size of large, comprehensive high schools (Lee & Ready, 2007). Proponents of small school reform argue that smaller high schools more often provide students with a rigorous curriculum and foster
environments in which students are more engaged, have higher attendance rates, feel safer, and graduate in higher numbers (Gates, 2005). The small schools movement attained a level of traction that is rare in education reform thanks to active state and district support and substantial bankrolling from multiple philanthropic organizations, most notably the Bill and Melinda Gates Foundation, which has poured roughly $2 billion into the creation of small high schools.

But the small school reform movement has not been implemented in a geographically neutral manner. Indeed, it is much more prevalent and well funded in urban school districts such as New York City, Chicago, Philadelphia, and Oakland. In New York City, where the reform has been implemented most aggressively, 27 large, comprehensive high schools were closed or downsized between 2002 and 2010 and replaced by more than 200 new small schools (Jennings & Pallas, 2010). The immense popularity of urban small schools movements was fueled by both a more general, prevailing sense of dissatisfaction with the academic intensity of high school coursework as well as the view that large, urban high schools, serving high proportions of minority and low-income students exemplify the worst qualities of American secondary education and are consequently prime targets for reform. In addition, the advantages of small schools were hypothesized to be particularly beneficial for upgrading the curricular experiences of disadvantaged student populations, who research suggests might otherwise fall through the cracks of large high schools, dropping out or skating by in low-level courses that do not prepare them for college or the workforce (Lee & Smith, 1997; Fowler & Walberg, 1991; Howley, Strange & Bickel, 2000; Bickel, Howley, Williams, & Glascock, 2001; McMillen, 2004).

Thus, in urban contexts, where schools and educators have struggled to meet accountability standards, small school reform was intended to support standards-based reforms by boosting the academic outcomes of some of the most disadvantaged student populations through improved academic rigor, increased personalization, and by utilizing community partnerships. And while
reform efforts to create small, urban schools are not entirely new and have a history dating back to the alternative school movement of the 1970s, the contemporary small school reform is distinct both in terms of its expanded national scope and its more explicit role as a turnaround strategy for large, underperforming urban high schools.

The rarely articulated but critical link between these coinciding factors – the call for “academic upgrading” and better college prep for disadvantaged students, the burgeoning urban small high schools movement, school closures, and the rise of the standards and accountability system – is a powerful assumption that shrinking urban high schools will necessarily lead to more disadvantaged students taking challenging courses and ultimately succeeding at higher levels in high school and beyond. Ironically, this alignment of reform and results remains empirically untested in a manner that can inform educators’ and policymakers’ understanding of the relationship between small school reform and its desired results. With billions of dollars and millions of children affected by small school movements nationwide, it is past time that researchers empirically interrogate these underlying assumptions—that small schools can improve disadvantaged students curricular experiences and academic outcomes.

Despite claims that small schools increase curricular rigor, no research to date has comprehensively evaluated such assertions. Disturbingly, while small schools in both New York City and Chicago have produced higher attendance and graduation rates, their record on improving student achievement has been less than convincing (Bloom et al., 2010; CCSR, 2010; CFE, 2010; Barrow et al., 2010). By 2008, even the most ardent supporters of small schools began to question the extent to which reducing school size alone can improve students’ college readiness (Gates, 2008). Because the extant research has not directly examined schools’ curriculum structures, the rigor of their courses, or their students’ course-taking patterns, we know little about the quality of
students’ academic experiences or the mechanisms through which small school reform improves (or fails to improve) student achievement.

This study fills the gap in the literature by directly addressing whether the small schools movement improved disadvantaged students’ access to a college preparatory curriculum or whether it simply repackaged and reintroduced historical inequalities in smaller bundles. Specifically, I focus on three elements of students’ curricular experiences. First, I examine schools’ curriculum structures, as defined by the number and type of courses offered and taken (Lee, Burkam, Chow-Hoy, Smerdon, & Geverdt, 1998). As discussed below, the hypothesized link between school size, the structure of the curriculum, and students’ course-taking patterns is crucial to policymakers’ assumptions that small school reform would produce more college- and career-ready graduates. Second, I assess how far students progress through the math curriculum. Specifically, I determine whether students in the newly created small high schools take more advanced courses in math than their counterparts in the traditional public high schools. Finally, I investigate whether the rigor of math courses in the new, small schools is equivalent to the rigor of courses in alternate high school options.

This study of small high schools, their curriculum structures, and their students’ progress through the math curriculum is especially important in light of evidence that political support for small schools was partially bolstered by research findings from national studies on school size with limited generalizability to the contemporary brand of urban small high school reform. As I will discuss in more detail in the next section, the intensely popular small schools movement and its transformation of secondary education in cities across the nation, was undertaken well before a research consensus on the effect of reducing school size could be established. My study, therefore, provides much-needed insight into the impact of the most recent attempt in the United States to adjust school size in order to accomplish important educational goals – an effort with a long history in American school reform.
Background

Limitations of Existing Research

As I will discuss in more detail in Chapter 2, the existing research on school size and students’ academic experiences can be divided into two eras:

1. Studies conducted during the late-1980s and 1990s, which focused on identifying the relationship between school size and educational outcomes among existing schools.

2. More recent studies conducted in the last 10 years, which have paid more attention to the results of contemporary small school reform (small by design).

These two eras of small school research are distinct in terms of both scope and methodology. Each provides significant contributions to our understanding of the academic implications of school size, but each also has shortcomings. The older body of work, much of which was conducted by Valerie Lee and colleagues, establishes the nation-wide relationships between high school size, schools’ curriculum structure, student course-taking, and academic achievement1 (Lee & Bryk, 1988, 1989; Lee & Smith, 1995; 1997; Lee, Croninger, & Smith, 1997; Lee et al., 1998). The most influential of these studies, conducted by Lee and Smith (1997), used Hierarchical Linear Modeling (HLM) and a nationally representative dataset (NELS: 88). Lee & Smith (1997) found that moderately small high schools of between 600 and 900 students had optimal results, both in terms of learning gains and the equity of learning within schools. Although the authors found that smaller

1 The research from the 1980s and 90s does not constitute the earliest research on small schools, but rather the earliest to empirically assess the academic consequences of school size (as opposed to the sociological consequences or economic efficiency). I discuss these two research strands briefly in Chapter Two, but for a more comprehensive review see Gladden, 1998; Lee & Burkam, 2003; Darling-Hammond, Ross, & Milliken, 2006.
schools (fewer than 600 students) were slightly more equitable, they also produced smaller average learning gains. Large high schools typically produced smaller learning gains and were much more internally inequitable than the schools in the 600-900 range. A consistent finding in the early research is that school size matters most for disadvantaged students, for whom large schools are particularly harmful (Lee & Smith, 1997; Fowler & Walberg, 1991; Howley, Strange & Bickel, 2000; Bickel, Howley, Williams, & Glascock, 2001; McMillen, 2004). These positive conclusions regarding smaller schools were influential in catalyzing the contemporary small high school reform.

A prominent theory for the positive academic outcomes associated with small schools was that they are more likely to have a “constrained” curricular structure consisting of largely college-prep academic courses and very few low-level courses (Lee & Bryk, 1988, 1989; Bryk, Lee, & Holland, 1993; Lee, Croninger & Smith, 1997). The constrained curriculum structure stands out in direct contrast to the large, “comprehensive” high school model, ubiquitous at the time these studies were conducted, which was characterized by a highly differentiated, “cafeteria style” curriculum in which students could select both which levels of courses to complete and how far to advance through the curriculum (Powell, Farrar, & Cohen, 1985; Angus & Mirel, 1999; Lee & Ready, 2009).

Indeed, the “constrained curriculum hypothesis” for explaining the positive associations between small school size and students’ academic outcomes was both informed and supported by decades of research on the negative consequences of this type of curricular differentiation for low-income and minority students, who disproportionately select (or are relegated to) lower-level courses and denied access to college preparatory knowledge when in large, comprehensive high schools. The theory was further buttressed by the empirical work on curriculum effects from the 1980s and 90s, which did not deal explicitly with school size, but linked differences in students’ academic course-

---

2 Note that contemporary small school reform such as the one in New York City have generally created schools in this smaller category, with enrollments of 550 students or less.
taking experiences to stratification in academic outcomes (Jones, Davenport, Bryson, Bekhuis, & Zwick, 1986; Garet & DeLany, 1988; Lee & Bryk, 1988, 1989; Lee, Burkam, Smerdon, Chow-Hoy, & Geverdt, 1998; Lee & Smith, 1995). Yet, because these studies on school size employed national data collected during the 1980s and early 1990s, they have limited applicability to contemporary urban small school reform for several reasons. First, research on small schools that utilizes national datasets relies on natural variation in school size (organically small schools) as opposed to a study of small schools that are part of an active small school reform (small by design) as is the case with the contemporary small schools movement. Second, research conducted during this era may have suffered from selection bias, resulting in inflated small school effect findings. Third, the contemporary small schools movement has created high schools that are substantially smaller (550 students or fewer) than the ideal size advocated by the early research. Finally, as I will discuss in Chapter 3, the standards and accountability movement, with its emphasis on upgrading the curriculum, has fundamentally changed the institutional context in which schools and students are making curricular decisions (Bishop & Mane, 1998; Rowan & Miskel, 1999; Fusarelli, 2002). This further limits the applicability of research conducted on school size and the curriculum during the 1980s and 90s to the contemporary reform.

The second body of research on small schools does not suffer from these same limitations. Because urban small school reform was already underway in many cities when this research was

---

3 Largely from two national datasets: High School & Beyond (HS&B) and the National Educational Longitudinal Study (NELS).

4 Put simply, the early research did not fully account for the possibility that small schools during this era may have been located in more socially cohesive contexts and served more motivated students than large schools (Stern & Wing, 2004). If small schools served students who would have taken more advanced courses and learned more regardless of what size school they attended, and the researchers did not fully account for this, then they may have overstated the relationship between small school size and these educational outcomes.
conducted, scholars were able to evaluate the effect of the reforms themselves, increasing the external validity of their results (see for example, Lee & Loeb, 2000; Wasley et al., 2000; Darling-Hammond, Ancess, & Ort, 2002; Bloom et al., 2011; Barrow et al., 2010; Schwartz et al., 2011). Furthermore, several of these studies employ sophisticated methodological designs, which address the selection bias concern (Bloom et al., 2011; Barrow et al., 2010; Schwartz et al., forthcoming). Generally, these studies find positive impacts of small schools on social engagement related outcomes such as student satisfaction, attendance, and graduation rates. However they do not find a consistent impact on student achievement or college readiness, particularly in mathematics.

A weakness of this more recent body of research, however, is its general lack of attention to the curriculum. Although these studies examine the effect of small schools on many educational outcomes including attendance and graduation rates, credit accumulation, and academic achievement, they do not evaluate the link between small school reform and schools’ curricular structures, the rigor of their courses, or students’ course-taking patterns. My dissertation seeks to complement and build on these existing studies by using sophisticated methodological techniques to shed light into the black box of schooling. By exploring student course-taking, this study addresses the mechanisms through which the small schools improve (or fail to improve) student achievement and will consequently help to make sense of the important findings discussed above.

**Contextualizing Contemporary Small High School Reform**

The recent small schools reform movement is not the first time that our nation has experimented with small schools. From the common schools of the 19th Century, small, locally grounded public schools have been a hallmark of the American educational system. In the 1970s, in a response to the perceived over-centralization of urban school systems and large, factory-modeled high schools, a progressive group of teachers and community organizations partnered to create a
few small, alternative, counter-cultural schools targeting students who might otherwise drop out of the system (Hemphill & Nauer, 2009). These schools were designed to be small so that teachers could get to know students better than in the large, anonymous high schools prevalent at the time. Grounded in their own pedagogical knowledge, the founders of the alternative small schools movement set out to create small, more personal school communities with stronger student-teacher relationships and more culturally relevant curriculum to better engage low-income and minority students and keep them in school. Interestingly enough, the epicenter of this movement was also New York City.

In some ways, contemporary small high school reform and this historical, alternative small school movement are alike: they both targeted a similar student population and they were both concerned with reducing inequality of educational opportunity. Yet these reforms were motivated by a different set of assumptions and goals regarding the potential effect of school size reduction. As discussed above, proponents of the contemporary wave of small school reform were undoubtedly influenced by both the standards and accountability movement and the early research on school size and thus were focused on boosting the academic achievement of disadvantaged students by reducing inequalities in access to a rigorous, college preparatory curriculum. In this way, the new generation of small school reformers’ expectations have transcended improvements in student engagement, and looked for higher achievement outcomes as well. Overlapping this outcome-oriented goal, many proponents of the contemporary reform movement were confident that small high schools could also produce more college and career ready graduates, even in challenging urban contexts. In a speech delivered at the National Education Summit on High Schools in 2005, Bill Gates explained:

Today, only one-third of our students graduate from high school ready for college, work, and citizenship. The other two-thirds, most of them low-income and minority students, are tracked into courses that won’t ever get them ready for college or prepare them for a family-wage job – no matter how well the students learn or the teachers teach. This isn’t an accident or a flaw in the system; it is the system. In district after district, wealthy white kids are taught Algebra II while low-income minority kids are taught to balance a checkbook! The
first group goes on to college and careers; the second group will struggle to make a living wage” (Gates, 2005).

Gates went on to propose a high school redesign initiative based on new principles—a re-vamped “Three R’s”—Rigor, Relevance, and Relationships. He then argued that these principles are almost always easier to promote in smaller high schools. This speech is illustrative of one of the implicit assumptions behind contemporary small school reform: the link between school size and the structure of the curriculum. While decrying the inadequacy of contemporary large urban high schools, Gates highlighted minority students’ disproportionate representation in low-track classes as a particularly egregious consequence of large, factory-model schools. The implication of his critique is that smaller schools would be less likely to sort students in this way, and that they will consequently improve disadvantaged students’ curricular experiences.

Importantly, this assumption—that small schools will provide disadvantaged students with a less differentiated and more academically rigorous curriculum than they would experience in large high schools—is the same “constrained curriculum” argument that was put forth by the 1980s and 90s researchers when seeking an explanation for their findings on the link between smaller school size and improved academic outcomes (Lee & Bryk, 1989; Lee, Croninger & Smith, 1997). Yet, as discussed above, it is unclear that the findings from the early research on school size will translate to the contemporary reform. Moreover, while the link between Catholic schools and a constrained, academic curriculum structure was well documented (Lee & Bryk, 1988, 1989; Bryk, Lee, & Holland, 1993) as was the link between a constrained curriculum and improved educational outcomes (Lee, Croninger & Smith, 1997), the direct link between small public schools and a constrained, rigorous curriculum structure was not proven. Thus, while we know with certainty that the new small schools created through the contemporary reform will not have the capacity to offer as many courses as large schools, we do not know if they will implement a constrained, rigorous, college preparatory
curriculum (constrained upwards) or if they will implement a constrained curriculum comprised of mostly lower-level courses (constrained downwards).

Put another way, while the 1980s and 90s research garnered much attention from educators, policy-makers, and philanthropic foundations desperate for a solution to the problem of chronically underperforming urban high schools, there are critical limitations in the applicability of this research to the current urban small school reform and its implications for disadvantaged students’ access to high status curriculum. Even Valerie Lee and her colleagues, authors of the most influential articles on school size from this era, caution that the small schools policy reform got out in front of the research base. They write, “The nation cannot expect that the generally perceived problems of contemporary U.S. high schools are going to be solved just by making the schools smaller….As important as school size is, our final policy recommendations focus on the curriculum” (Lee, Smerdon, Alfred-Liro, & Brown, 2000, p.165).

More recently, critics have questioned whether urban small school reform is a strong enough policy action to succeed at the difficult task of reducing educational inequalities in our country’s most challenging, and highly stratified urban contexts. For example, outspoken opponent of segregation, Jonathan Kozol proclaimed:

Small schools are usually less chaotic than big schools; they are sometimes more intimate and relaxed than big schools. But the small school concept, which no one is proposing for the schools in white suburban districts, is essentially an anti-riot strategy for segregated children, an anti-turbulence measure, a short-term solution to perceived chaos in large segregated schools. Small, segregated, and unequal schools are only an incremental improvement over large, segregated and unequal schools. They don't address the basic issues (Kozol, 2006).

This critique—that small school reform is a band-aid solution that will not fundamentally improve disadvantaged students’ educational outcomes—is far from the only complaint. Some educators have raised concerns that target the curriculum itself. They caution that small schools may not be as effective for all student subgroups—for example, their limited capacity to offer a wide
range of courses may inhibit initially high-achieving students from accessing the same range of advanced courses that might be available to them at a larger school (Lee & Ready, 2007; Hemphill & Nauer, 2009; La Quijota, 2009). Others fear that very small schools (those with fewer than 600 students), particularly those serving predominantly disadvantaged students, may offer only lower-level courses and thus deny their students access to a college preparatory curriculum (Hemphill & Nauer, 2009). My dissertation adds to the existing literature on small school reform by empirically testing whether these fears have been born out or whether the new small high schools implemented the type of rigorous, college preparatory curriculum that the proponents promised.

It is also important to recognize that small school reform is being asked to improve educational equity in challenging, deeply inequitable urban contexts. For example, from the New York City high school class of 2009, many of whom attended small schools, only 13 percent of black and Hispanic students graduated with Regents Exam scores that designated them as “college or career ready” (Otterman, 2011). By contrast, half of all Asian students and over 40 percent of white students reached these standards.

At its heart, reducing school size is an organizational strategy, which is being utilized in part to address this type of inequality. Especially in urban environments like New York City, small-schools movements can be viewed as bold efforts to leverage the beneficial organizational properties associated with smaller institutions—less anonymity, less structural differentiation, improved member engagement—to help those students at the bottom of the stratified education system. And while adjusting school size may seem like a blunt instrument for fostering equity and leveling access to a rigorous curriculum, this is the type of tool that education policy-makers have at their disposal. In our current, post-civil rights era, in which strategies like desegregation are no longer politically viable, education reformers are increasingly looking to policies that can aid in producing equitable academic experiences for all students without creating integrated schools (Darling-Hammond, 2010).
But were proponents of urban small school reform right to assume that breaking up large, highly differentiated high schools into smaller institutions would reduce inequalities in the educational experiences of advantaged and disadvantaged students? Is school size indeed the critical variable in this equation, such that simply creating smaller schools will improve low-income and minority students’ access to high-status knowledge? This dissertation will address these questions by examining the impact that small high school reform has had on student course-taking and curricular rigor in the New York City public school system.

Clarification of Terms

In the literature on the curriculum, the concepts of *upgrading the curriculum* or *academic intensity* typically refer to two aspects of the course work that students take in high school: the breadth of study and the intellectual level of difficulty. An intensive course of study generally implies that students take many courses covering a range of *academic* subjects such as English, Mathematics, Science, History, and foreign languages *and* that these courses are advanced with regards to their content (for example, taking Calculus instead of Pre-Algebra; Attewell & Domina, 2008). The concept of curricular *rigor* has sometimes been used interchangeably with the concept of intensity, however throughout this study, I use the concept of rigor to refer to teachers’ standards of performance within courses (Adelman, 1999). In other words, Algebra I can be taught in an academically challenging, or rigorous, fashion where students are held to high standards of content mastery or it can be watered down in such a way that students can pass the course without learning the appropriate skills.

Finally, the academic intensity of students’ curricular experiences is delimited by their school’s *curriculum structure*, as defined by the number and type of courses offered and taken (Lee, Burkam, Chow-Hoy, Smerdon, & Geverdt, 1998). For example, some schools may implement a
differentiated curricular structure, characterized by a wide range of course offerings at varying levels of difficulty, while others offer a constrained curriculum, characterized by a narrow band of largely academic courses. As discussed above, the hypothesized link between small school size, a constrained, academically intense, and rigorous curriculum, and student course-taking is crucial to policymakers assumptions that small school reform would produce more college and career ready graduates.

**Focusing on Math**

In order to evaluate the relationship between school size and student course-taking, this study will focus on students’ experiences in the math curriculum. Math is singled out for both methodological and substantive reasons. Substantively, evaluating the effect of small school reform on students’ math experiences is critical for policy and equity reasons. The federal government has made improving science, technology, engineering and math (STEM) education a particular priority, asserting that the nation’s future economic health depends on building a strong workforce in these fields (Office of the Press Secretary, 2010; Adams, 2014).

Thus, in today’s technology-based economy, persistence in the high school math curriculum provides students with the academic credentials and tangible skills needed for success in both college and the workforce, which are increasingly seen as requiring an equivalent level of knowledge and skills (ACT, 2006). Advanced high school math courses also make students more attractive to four year colleges (particularly the most selective colleges), reduce the need for remediation once in college, and prepare students for higher education trajectories into sectors of the labor market characterized by growth, stability, and above-average wages (Adelman 2006, National Academy of Sciences, 2007; Saldler and Tai, 2007; Langdon, McKittrick, Beede, Kahn, & Doms, 2011; Crosnoe & Schneider, 2010). While the once pervasive gender gap in math coursework has closed (Freeman,
Black, Hispanic, and low-income students continue to persist less far in math during high school (Crosnoe & Schneider, 2010; Crosnoe & Huston, 2007; Lucas, 1999; Oakes, 1990) and are less likely to meet college readiness benchmarks in this subject than whites or Asians (ACT, 2011). Furthermore, a recent report issued by the government finds that Black and Hispanic workers are severely underrepresented in lucrative STEM jobs (Beede, Julian, Lehrman, McKittrick, Langdon, & Doms, 2011).

Moreover, it has become increasingly clear that it is not just the number of math credits, but which courses students take that matters. In other words, three years of remedial math do not produce the same positive academic and post-secondary outcomes as three years comprised of Geometry, Algebra II, and Pre-Calculus. Research indicates that students who progress farther through the math pipeline (i.e. reaching more advanced courses) learn more math during high school and have better odds of attaining a bachelor’s degree (Alexander & Pallas, 1984; Adelman, 1999, 2006; Lee & Bryk, 1988, 1989; Lee et al., 1998; Gamoran & Hannigan, 2001; Natirello, Pallas, & Alexander, 1989; Horn, Kojaku, & Carroll, 2001). Specifically, progressing past Algebra II seems to be a critical determinant of post-secondary participation and success (Adelman, 2006). Analyses conducted by the ACT found that among urban ACT takers more than half of students persisting through the math pipeline to Calculus, but only seven percent of those persisting only to Geometry, meet the benchmark for college and career readiness in Math. Thus, if small schools can succeed at the task of getting disadvantaged students further through the math sequence, they may help close the racial and socioeconomic gaps in college readiness and access and remove some of the barriers

---

5 ACT test-takers living in one of the 66 large urban districts that are members of the Great City Schools Coalition.

6 It is important to note that these “curriculum effects” studies may also suffer from selection bias. This possibility will be discussed in greater detail in Chapter 3.
to equal participation in lucrative STEM careers. This study will be the first to interrogate whether small school reform can help achieve these important goals.

Furthermore, math is ideal for testing whether small schools implement a more constrained, academically intense curriculum, because math has typically been among the most highly differentiated subjects in comprehensive high schools, both in terms of level (remedial, regular, honors, AP, etc.) and in terms of which courses are offered and taken (Algebra, Geometry, Trigonometry, Calculus, etc.; Lee, Croninger, & Smith, 1997). The potential complexity of the math curricula allows for substantial variation in students’ experiences in these subjects both within and between schools. In addition, extant research suggests that math is an area of the curriculum where learning is particularly responsive to school experiences, making it ideal for evaluating the impact of a school attribute such as size (Murnane, 1975). Moreover, studies on the outcomes of tracking and ability grouping find that these practices have the largest and most consistent effects in math (Gamoran, 1987; Lee & Bryk, 1988).

Math also has, in theory, a clear linear sequencing of courses where skills and concepts build on one another throughout the sequence, meaning that curriculum in one class must generally be mastered before students can successfully move to the next course in the sequence. Methodologically, this allows for a clear and elegant measure of students’ educational attainment in math—simply the most advanced level of course taken during high school. Finally, math courses are more easily identifiable by course title than most other subjects. For example, “Pre-Calculus” has more specified content than “English II” (Lee, Croninger, & Smith, 1997). Yet even in math, there is a legitimate concern as to whether course titles accurately represent the content and difficulty of what is actually taught and learned.
New York City—The Ideal Location for a Study of Small High School Reform

To evaluate the impact of small high school reform on students’ experiences in the math curriculum I employ a unique, longitudinal dataset with full transcript files for all students in the New York City public school system. New York City is the ideal location for evaluating the potential of small school reform because it has adopted this policy more aggressively than any other school system in the country. And because New York State mandates extensive high school exit exams, known as Regents exams, I will be able to go beyond course labels and actually assess curricular rigor.

Finally, as the largest school district in the nation, New York City Public Schools have more students currently enrolled in small high schools than are enrolled in all of the high schools in the country’s 7th largest school district, the Houston Public Schools (Bloom et al., 2010). The sheer size of the New York City school system allows me to conduct a large-scale evaluation of small high school reform within one district. This is important because unlike studies using national datasets to evaluate the effect of school size, focusing on one district allows me to remove all across-district, unmeasured policy differences. Thus, my dissertation is the first study to bring in-depth analysis of student transcript data to bear on the evaluation of urban small school reform in the city that has gone furthest to implement it. Moreover, the New York City public school system’s size and racial and economic diversity allow my findings to translate to urban contexts far beyond the five boroughs.

Course Titles Versus Course Content

When researching the impact of school size on students’ curricular experiences it is important to heed the caution of institutional theorists, who believe in the necessity of evaluating the extent to which educational reforms reach the technical core of schools. In other words, how do we
know that a course designated as “Algebra II” is actually teaching students the appropriate content for that subject? Should we believe that students receiving an “A” grade in Calculus in one school necessarily have achieved the same level of mastery as students receiving equally stellar grades in Calculus in another school? Institutional theory suggests that schools and teachers may feel pressure to follow a “common script,” leading them to make decisions based on widely held institutional rules and beliefs of what “real school” looks like, particularly when it comes to curricular decisions (Metz, 1989; Tyack & Cuban, 1995; Riehl, Pallas, & Natriello, 1999).

As state and federal pressure to “upgrade” the curriculum becomes more and more intense, and as the belief that all students should follow a college preparatory curriculum becomes institutionalized, all schools are likely to feel pressure to project the image of a strong academic curriculum, regardless of whether they have one or not. And schools with substantial low-income and minority student populations may feel this pressure even more intensely than their more advantaged counterparts. Not only are they less likely to offer strong academic curricula in the first place, they are also more likely to enroll students who lack sufficient preparation for higher-level coursework. Indeed, “course credit inflation” is becoming an increasing problem (Dougherty, Mellor, & Jian, 2006a). Similar to the concept of grade inflation, in which the knowledge and skill of the average student receiving an “A” declines over time, the level of content mastery for the average student receiving credit for a course with a given title may be declining over time.

Low-achieving NYC high schools may experience the pressure to engage in course credit inflation more acutely than comparable schools elsewhere since they enroll many students who are severely behind grade-level and are operating in a state where students must pass high-stakes Regents exams in order to graduate. Increasingly, schools realize that their students must take core academic courses in order to graduate, however the lack of students passing the Regents exams with
scores that indicate “college readiness” suggest that curricular rigor may not have increased, or at least not in a way that has translated to improved outcomes.

In short, it is increasingly important for research on secondary education not to blithely assume that course titles are accurate representations of course content or that students who have received credit for a course have necessarily demonstrated mastery over its subject matter. I will address this concern in the small schools context by examining the relationship between students’ performance in a course (final grade) with their performance on the NY State Regents Exam for the relevant subject. I will also assess the extent to which the link between succeeding in a course and succeeding on the relevant Regents Exam varies as a function of school size. Thus, this study will be the first to assess whether “course credit inflation” occurs in New York City high schools and, if so, whether it is related to school size.

Organization of the Dissertation

In the next chapter, I will argue that the early, positive research—the first era of research I discuss above from the 1980s and 90s—linking small schools to improved academic outcomes has serious shortcomings when it comes to predicting the results of the contemporary reform. Furthermore, I explain that the more recent research, which does focus on the contemporary reform, has failed to find consistent positive findings regarding student achievement and has not yet evaluated student course-taking or curricular rigor. Following this review of the research most tightly focused on my topic, I will also review a broader body of research on the historical evolution of the relationship between the curriculum, school size, and educational equity in the United States and discuss the implications of this history for the current reform.

In Chapter 3, I will discuss the existing literature on the relationship between students’ course-taking and their educational outcomes such as learning during high school and post-secondary success. I will then outline a conceptual framework for understanding the relationship
between urban small school reform, curricular structure, and students’ access to high status knowledge and I will discuss the ways in which the current institutional context may shape schools’ and students’ decisions regarding the curriculum. In this framework, the curriculum in each subject is conceptualized as a sequence of educational opportunity, where a student’s opportunity to learn advanced content is partially dependent on their successful completion of the preceding courses in the sequence (Stevenson, Schiller, & Schneider, 1994). How far students progress through these curricular sequences, or “pipelines,” is determined by two interacting forces; first, the structure of the curriculum—both what’s offered and how students are assigned to courses; and second, students’ own decisions, motivation, ability and performance (Sorensen & Hallinan, 1977). Importantly, the exact same student might experience vastly different educational experiences in different schools with different curricular structures (Sorensen, 1987). Consistent with the “constrained curriculum” hypothesis, school size is conceptualized as shaping the structure of the curriculum—particularly its level of differentiation—yet other mediating factors such as student body composition, teacher quality, and the institutional context are hypothesized to influence schools’ curricula as well.

In Chapter 4, I discuss how my conceptual framework shapes the methodological design of my study. Of particular importance is the need to account for selection bias when comparing students in the new small high schools to their counterparts in the alternate public high school options in New York City. To accomplish this I utilize multiple propensity score matching techniques to select appropriate control groups for comparison to students attending small high schools. Next, in Chapter 5, I present my descriptive results, which situate the new small high schools and the students they enroll within the broader landscape of high school options in New York City. In particular, I illustrate the ways in which high schools in New York City are highly stratified both in terms of student demographics and access to advanced mathematics courses.
In Chapter 6, I present the findings from my more complex descriptive analyses delineating the associations between school factors such as size and selectivity and students’ progress through the math curriculum, controlling for students’ socio-demographic and academic characteristics. I also test whether attending a new small high school may be more beneficial for disadvantaged students, as the extant literature suggests. Next, I present my analytic estimates of the effect of attending a new small high school on students’ progress through the math curriculum. Finally, I assess the rigor of courses across high schools in New York City by estimating the gap between students’ grades in specific math courses and their marks on the Regents exams covering those topics. Specifically, I explore whether students’ performance in their courses and their performance on the relevant exams is more tightly aligned in some schools than others and, in particular, whether this alignment varies as a function of school size. Finally, I conclude this dissertation in Chapter 7 with a set of conclusions drawn from my analysis of the relationship between small high school reform and students’ experience in the math curriculum. I also provide a set of realistic policy recommendations and I reflect on the limitations of this study and avenues for future research. I turn now to the specific research questions around which my dissertation is organized.

**Research Questions**

This study builds on the existing literature on school size and the curriculum by investigating the relationship between the current small school reform in New York City and students’ course-taking experiences in math. Employing a unique longitudinal data set with complete student transcript data, I will utilize a propensity score matching approach within a multilevel modeling framework to address the following questions:

1. What differences exist between the curriculum structures of small versus large high schools in New York City?
2. Do students in the new, small, high schools take more advanced courses in math than their counterparts in the alternate New York City public high school options?

   a. Does the relationship between attending a new, small school and student course-taking depend on students’ socio-demographic characteristics or initial achievement levels?

3. In terms of academic rigor, are courses in New York City’s new, small high schools comparable to those across the range of alternate public high school options?

**Significance**

Despite the popularity of urban small high school reform and the strong policy focus on this issue, no study to date has comprehensively examined how the recent shift from large to small high schools has influenced the rigor and equity of students’ course-taking patterns. A new, sophisticated, empirical investigation of this issue is vital considering that the early research base may have limited applicability to the contemporary reform and the more recent research has not examined the curriculum. Given the importance of students’ academic experiences for their successful transition to higher education, understanding the impact of small school reform on student course-taking will be a critical step in the broader evaluation of small school reform.

In addition to shedding light on this crucial policy issue, this study also contributes to our broader understanding of how an organizational feature of schools—this size—impacts students’ educational opportunities. Furthermore, by focusing on disadvantaged students’ experiences in the math curriculum, this study augments our knowledge of the factors influencing the racial and socio-economic gaps in STEM education and careers. Finally, this study assesses what students actually learn in their courses rather than taking course grades and course labels at face value.
Chapter Two

Review of Relevant Literature

In developing a study on the relationship between small school reform and student course-taking, I wanted to place my research within the existing literature on school size and students’ access to high-status curriculum. This required me to examine several somewhat distinct, and often disconnected bodies of literature, including research on school size and research on the curriculum. In this chapter I will highlight the studies from both strands of research that most directly inform my understanding of the association between small school reform and students’ curricular experiences in high school. In particular, I will focus on the early research, conducted during the 1980s and 90s, that first established the positive link between small school size and students’ academic outcomes (and the shortcomings of applying it to current small school reform). Then I discuss the more recent era of research evaluating contemporary urban small school reform policies, emphasizing both what these studies have revealed as well as what we still do not know.

Next, I situate my analyses within an understanding of the historical evolution of the link between the curriculum, school size, and educational equity in the United States. I then review the history of small school reform in New York City to provide context for my research and to offer a concrete touchstone from which to assess the strengths and weaknesses of the relevant literature. This review, which combines both literature and history, has helped me design my study both to build on the existing research base and to remain grounded in New York City, while at the same time considering applications to small school reform in other contexts. The historical overview highlights school characteristics in addition to school size, such as student body composition, as important in determining students’ access to high-status curriculum. It has also led me to question whether reformers were correct to assume that the increased curricular constraint associated with reducing school size would necessarily manifest as “constraining upwards” to a rigorous, college
preparatory level of difficulty and to consider potential mediating factors that might influence the relationship between reducing school size, the level at which schools will constrain their curriculum, and students’ experiences in math. Finally, I save a discussion of some of the research on curriculum effects for Chapter 3 because it pertains directly to my conceptual framework.

School Size Research

The Early 20th Century Research—Building the Case for Small Schools

In this section of my literature review, I will focus on reviewing and critiquing the academic research on school size and students’ academic outcomes from the 1980s and 90s because this was the body of literature that was most influential in fueling the contemporary small high school reform and in developing the constrained curriculum hypothesis. But before delving into these studies, it is important to acknowledge that there was research on school size prior to the advent of widespread achievement tests that did not focus on students’ academic outcomes. Research from this era focused on issues pertaining to economic efficiency and social relations within schools (Lee, 2000). Since the mid-1990s however, the majority of research on school size has focused on its impact on academic, outcome-oriented indicators such as student achievement (both eras of research discussed in Chapter 1 fall into this category). Thus, I will briefly summarize the sociological and economic research strands on schools size before focusing in more depth on the research linking school size to academic outcomes (for a more comprehensive review of the sociological and economic research see Gladden, 1998; Lee & Burkam, 2003; Darling-Hammond, Ross, & Milliken, 2006).

The Sociological Argument for Small Schools. Organizational theorists have long recognized that size is an important ecological feature of an organization, affecting both its structure and social dynamics. As an organization grows, human interactions and ties become increasingly hierarchical and formal as bureaucratic structures replace human relationships (Weber, 1947). The
internal structures of organizations also tend to get more complex and differentiated as they increase in size.

Beginning in the 1960s, prior to the accountability movement and proliferation of test score data, sociologists of education turned this organizational lens upon schools, concluding that as schools grow larger, they become more anonymous and meaningful personal interactions and relationships among teachers and students become more difficult to maintain (Barker & Gump, 1964; Bidwell, 1965; Bryk & Driscoll, 1988; Bryk, Lee & Holland, 1993; Sizer, 1984). Conversely in small schools, there are more opportunities for both formal and informal interactions between teachers and students. As a result, teachers are more likely to get to know their students on a non-superficial basis and students are consequently less likely to “fall through the cracks” (Raywid, 1996). “Marginal” students are particularly affected by school size in terms of participation and engagement (Barker & Gump, 1964). The academic research on small schools from this sociological strand of research is reasonably consistent: small schools perform better on outcomes that are social or affective in nature such as isolation, alienation, or social engagement (Chambers, 1981; Fowler & Walberg, 1991; Newmann, 1981). Studies have also found that students in small schools are more likely to participate in school activities, are more satisfied with school, have lower dropout rates, and have higher attendance rates (Lindsay, 1982; Pittman & Haughwout, 1987; Gladden, 1998; Lee & Burkam, 2003).

Increased personalization and student engagement does not directly influence the structure or quality of the curriculum that students will experience, yet it may indirectly impact the degree to which students are able to profit from the curriculum that is in place. As Sorensen and Hallinan (1977) explain, there are two components to the educational process: what students are taught and how much they learn of what they are taught. Opportunities to learn interact with the effort and ability of students to produce learning. Thus, while improved student engagement may not affect the
structure of the curriculum, it might influence how much students learn in what courses they do take. Furthermore, if students are more motivated and have better attendance in smaller schools, they may fail fewer courses and be more inclined to take on courses beyond what is mandated by graduation requirements. If either of these phenomena occurs, disadvantaged students in smaller schools may be expected to progress further through curricular pipelines than their counterparts in large schools, within the constraints of their schools’ curricular offerings. Clearly the structure and quality of the curriculum are still crucial; no matter how motivated the student, they can neither take courses that are not offered nor learn challenging content if only remedial levels are available.

Finally, organizational research also confirmed that as average school size increases, so too does the number of course offerings (Haller, Monk, Spotted Bear, Griffith & Moss, 1990; Monk & Haller, 1993). Small schools, on the other hand must focus their resources on a core set of courses and programs, absorbing students at both ends of the distribution of ability or interest into courses that they might not otherwise have chosen (Monk, 1987; Monk & Haller, 1993). These findings lend support to the constrained curriculum hypothesis, and they make sense intuitively—small schools simply lack the capacity to offer a differentiated curriculum. In my conceptual framework, I will argue that this can be good or bad news depending on complex interactions between students’ achievement level when they enter a small school, the composition of the student body, and how the curriculum is constructed to meet not only mandates but also the perceived academic needs of the majority of students. Thus, this early sociological research is important to my understanding of the role of school size in shaping students educational experiences, however, this older generation of studies did not delve deeply into the content or rigor of schools’ course-offerings so it is difficult to diagnose the relationship between school size and the academic intensity of students’ coursework.

The Economic Argument Against Small Schools. Traditionally, economists have viewed schools as no different from any other service-production organization; increasing the number of
students can increase efficiency of the organization under two conditions (Buzacott, 1982; Lee, 2000). The first is simply if the cost of supplies is reduced by purchasing in bulk. The second condition is if increasing the number of service recipients maximizes the efficient delivery of that service. For example, if schools plan on tailoring their curriculum to the ability levels of their students, then more students can help maximize the efficient delivery of instruction by increasing the number of students with similar ability. In other words, it’s more efficient to have an AP chemistry class with 20 students than one with three. According to this logic, large schools can more efficiently offer an expansive and differentiated curriculum (Kenny, 1982). Thus, the advocates of the economy-of-scale argument generally believe that curricular differentiation is beneficial because it more efficiently targets academic courses to students of varying abilities and interests. As I will discuss in greater detail below, this belief in the value of curricular differentiation was dominant for the majority of the 20th century in American education, however it is fundamentally opposed to the contemporary belief that all students should have access to a college preparatory curriculum.

This argument rests on an assumption that greater size results in greater economic efficiency and will consequently produce savings. Other research on this topic has called into question whether this assumption holds for educational organizations, which tend to increase their administrative staffs as they grow in size (Chambers, 1981; Fox, 1981). Recent studies conducted on small schools in New York City find mixed results. One study concludes that small schools in NYC have significantly higher per-pupil expenditures than their larger counterparts (Jennings & Pallas, 2010). Other studies have found that NYC small schools generally have equivalent operating costs and lower costs per graduate (Public Educational Association, 1992; Steifel, Berne, Iatrola, & Fruchter, 2000) but that cost structures differ across the schools depending on their “mission”— comprehensive or themed (Stiefel, Schwartz, Iatrola & Chellman, 2009). Themed schools minimize
per pupil costs at smaller enrollments than comprehensive high schools (Stiefel, Schwartz, Iatrola & Chellman, 2009).

However, it is important to note that in contrast to the economic efficiency argument in favor of large, comprehensive high schools, small school advocates have never framed their reform in terms of economic efficiency. Instead, small schools have been promoted as strategies for improving educational equity and raising academic outcomes, especially for disadvantaged students. Thus, I turn now to the research on school size and small school reform that directly speaks to these issues.

**Early Research Linking School Size to Academic Outcomes.** I begin my discussion of the literature on school size and students’ academic outcomes by summarizing the first era of research discussed in Chapter 1, conducted during the 1980s and 90s, which predominantly utilized national datasets. Research more systematically focused on the connection between school size and academic outcomes developed partially out of research on Catholic schools. Researchers found that Catholic schools, which are smaller, on average, than public schools, were producing better academic results, even after controlling for student demographic characteristics. They found that students were learning more in Catholic schools and that the link between students’ socio-demographic characteristics (such as their race/ethnicity or social class) and their learning was weaker in Catholic schools, meaning that they were also more internally equitable (Lee & Bryk, 1989; Bryk, Lee, & Holland, 1993). A prominent theory for the Catholic school advantage was curricular structure—Catholic schools are more likely to have a “constrained” academic curriculum consisting of largely academic courses and very few low-level courses whereas the contemporary public high schools typically offered a broader, “comprehensive” curriculum (Bryk, Lee, & Holland, 1993; Lee & Bryk, 1989). This was the origin of the constrained curriculum hypothesis.
The Catholic school curricular advantage was soon confirmed—researchers found that students take more advanced math courses in Catholic school, even after controlling for entering achievement, and that course-taking is far more equitable in Catholic as compared to public schools (Lee et al., 1998). The small size of Catholic schools was thought to be one of the factors, alongside their social justice mission, leading them to offer a constrained academic curriculum and thus inducing their students to take more advanced courses and ultimately learn more during high school.

Motivated by the promising findings on Catholic schools, research began to expand beyond sector differences to explicitly tie school size to student outcomes (Lee & Smith, 1995, 1997; Lee & Loeb, 2000; Wasley, Fine, Gladden, Holland, King, Mosak & Powell, 2000). The results from this series of studies were similarly positive: students attending smaller schools learned more and learning was more equally distributed by race and socioeconomic status. The most influential of these studies was conducted by Valerie Lee and Julia Smith (1997). Using Hierarchical Linear Modeling (HLM) and a nationally representative dataset (NELS: 88), Lee & Smith found that moderately small high schools of between 600 and 900 students had optimal results, both in terms of learning gains and equity. Although the authors found that smaller schools (fewer than 600 students) were slightly more equitable, they also produced smaller learning gains. Large high schools, on average, both produced smaller learning gains and were much more inequitable than the schools in the 600-900 range. However, Valerie Lee and colleagues doubted that school size itself could produce these dramatic results and consequently focused on schools’ curricular structures as the mediating factor between school size and student learning in future work (see section on curriculum effects in Chapter 3).

A consistent finding in this early research on school size is that it impacts the educational outcomes of disadvantaged students most strongly, with large schools being particularly harmful for this demographic (Lee & Smith, 1997; Fowler & Walberg, 1991; Howley, Strange & Bickel, 2000;
Bickel, Howley, Williams, & Glascock, 2001; McMillen, 2004). The fact that research studies on school size during this era consistently found that small schools were especially beneficial for disadvantaged students’ academic outcomes further legitimized it as a promising reform strategy for under-performing urban high schools.

Yet given the many limitations of these findings, which I discuss below, we have to question how generalizable they are to the contemporary, urban small high school reform. In the next section of this literature review, therefore, I will address this question by dissecting the limitations of the early academic research on school size for supporting the current movement. I will then review the more recent research on contemporary small school reform.

**Limitations of Early Research for Supporting the Current Movement**

By the turn of the 21st century, a convincing body of research, outlined above, had accumulated suggesting that students, particularly those from disadvantaged backgrounds, benefit academically from attending smaller high schools. At a time when educators, policymakers, and philanthropies were prioritizing the development of strategies to remedy the problem of “failing” urban high schools, the confluence of positive research on small schools did not go unnoticed and spurred investment in substantial small school creation in many cities around the nation. But now that hundreds of small schools have opened and billions of dollars have been spent, doubts are beginning to be raised as to whether small schools really are the silver bullet solution that advocates forecasted and the academic research seemed to support. These doubts may be grounded. For a number of reasons, the late 20th century research base is limited in its applicability to large-scale urban small schools policy reform.

First, virtually all research from this era on the links between school size and academic outcomes suffered from some degree of selection bias. Second, the positive findings that emerged from the early research base, particularly those studies using national samples, may not generalize to
the contemporary policy reform, particularly in the context of large, urban school systems characterized by bureaucratic regulation, and intense between-school stratification by race, class, and academic achievement. Third, the current reform intentionally created smaller schools than recommended by the extant literature. Finally, results from research conducted in the 1980s and 90s may not translate to the current institutional context in which secondary schools experience increased institutional pressure to guarantee that all students receive a college preparatory education.

**Selection bias.** The majority of the early research on school size and academic outcomes that utilized large-scale state or national databases compared existing small schools to existing large schools. However, schools that were already small or mid-sized in the 1980s and 90s, despite the dominance of the large, comprehensive high school model during that era, may have also been different in other, unmeasured ways (Stern & Wing, 2004). For example, it is possible that small, rural or mid-sized, suburban schools tend to be located in more cohesive, supportive, and homogenous communities and large urban schools tend to be located in areas that have more concentrated challenges to academic success. As discussed above, some of the early influential research on schools size included Catholic schools, which are usually small but are also different on a host of other attributes such as student advantage, parental motivation, and school mission. While robust statistical controls for student characteristics, urbanicity, sector, and school funding help account for these potential external forces, they may not capture everything. These unmeasured community and student body composition differences could have affected both how schools structure their curriculum and/or students’ academic outcomes rather than differences in school size.

A smaller scale, field-based study on school size conducted by Lee, Smerdon, Alfeld-Liro, & Brown (1996) illustrates the potential limitations of the national studies. The authors varied their sample selection both on urbanicity and sector (a small and large urban, suburban, and rural public
school, and a small and large Catholic high school). In terms of the curriculum, both the small and large Catholic high schools followed the typical constrained, academic curriculum. Interestingly, the curriculum in the small public high schools did not follow the Catholic school model. Instead, teachers in the small public schools tried hard to offer the same type of curriculum as a large comprehensive high school, but failed due to a lack of capacity. They held to the belief that a differentiated curriculum was the best way to serve the diverse needs of their student body. The combination of this desire to differentiate with a small organization structure resulted in “unusual and illogical” course sequences. For example, each year in the small rural school, faculty would compile a list of courses they could offer and students would vote on which they wanted. As a result, students might begin a foreign language (such as Spanish) but never get a follow-up course. Many seniors took physics but few took advanced math courses simply because they disliked the math teacher (there was only one). Indeed, the authors found that the majority of small schools in their study did not want to be so small. While limited in scope, this study is illustrative of the type of shortcomings that can arise when attempting to apply research findings to a new context.

Teachers in new small schools may be self-selected as well (Stern & Wing, 2004). If teachers who are more motivated, committed, energetic, or creative choose to work in small schools, then positive results generated by these schools may not generalize to the system as a whole. Stern & Wing (2004) caution that attempts to reduce the size of all high schools may be based on a “fallacy of composition,” a mistaken hope that what is observed in specific cases can be generalized to the whole high school population. Finally, selection bias can create problems at the school-level as well. Schools may try to meet the needs (perceived or actual) of their students through the curriculum. Consequently, they may respond differently to different types of student bodies in terms of which courses they offer. Furthermore, socio-economically advantaged parents may be more likely to lobby successfully for more advanced courses. Consequently, if selection into smaller schools results in a
more advantaged, academically prepared, or motivated student body, studies that do not adequately address selection may overstate the link between small school size and the provision of a “constrained” academic curriculum.

**Not all small schools are created equal.** Beyond methodological limitations, there is a related concern that the early research on school size, particularly the studies based on national samples, might not be generalizable to the creation of new small schools. The small schools captured in national datasets during the 1980s and 90s were likely small for a host of different reasons unknowable to the researchers. Some were likely small by default, particularly rural schools, because of the size of the community they served. Others may have been small because it was deemed conducive to their mission—whether private, Catholic, or themed magnet. It is even possible that some schools from the “alternative” small school movement of the 1970s were sampled in these datasets. So while some of the sampled schools may have been similar in design to the schools created by the contemporary reform, most were likely small for other reasons. Furthermore, because researchers had no way of knowing which schools in these datasets were small for which reasons they could not disentangle the potentially varying effects of these different types of small schools.

The current policy reform creates a very specific type of small school, targeted towards specific students, with specific goals in mind. It takes a mechanical approach to small school creation, in some cases closing chronically underperforming large high schools and opening smaller schools—schools that meet specific district requirements, selected through a competitive request-for-proposal application process—to serve as replacements. This is a fundamentally different type of school from a rural school, which is small by default or a Catholic school which may be intentionally small, but also differs on a host of other characteristics ranging from student selection to school mission. Thus the findings generated from one type of small school may not accurately predict the educational outcomes of the other.
Moreover, current small school reform has been principally adopted in large urban school districts such as New York City, Chicago, Philadelphia, and Oakland, where new small schools have been targeted to serve the most disadvantaged students. Because it has been implemented predominantly in historically underperforming contexts, the new small schools movement will experience challenges that the average school does not. For example, based on 2011 results from the National Assessment of Educational Progress (NAEP), 41% of all New York City eighth graders have “below basic” math skills. And of course, this masks huge sub-group differences—half of all black and Hispanic students rated below basic, while only one fifth of white and 14% of Asian students did so. Moreover, over 40% percent of students receiving free lunch, and over 80% of English Language Learners were rated as below basic. In schools where the majority of entering students are performing below grade-level, it may be particularly difficult to effectively implement a rigorous, college preparatory curriculum in a way that is still tailored to students’ needs. If school size is hypothesized to affect student learning primarily through the structure and focus of the curriculum, and small schools are being targeted to serve predominantly disadvantaged and underperforming students, then it is important to understand the link between a school’s student body and its curricular structure (see conceptual framework).

**Contemporary small high schools may be too small.** One weakness of the literature on small schools, particularly the early research, is a lack of consistency regarding what actually constitutes a “small school” (Darling-Hammond, Ross, & Milliken, 2006). However, the most influential studies from this early era of research—those conducted by Valerie Lee and colleagues—identified high schools of 600-900 students as the ideal size for maximizing both equity and excellence (as measured by learning gains during high school). This is substantially larger than the average size of schools introduced during the current small high school reform. As I will discuss in greater detail below, the new small high schools in New York City were intentionally created to
enroll 550 students or fewer. According to Lee and Smith (1997), this size of school should produce the most within-school equity of outcomes, but should also generate lower learning gains than schools of 600-900 students. In short, the early positive research findings on small schools may not translate to the contemporary reform simply because the reformers did not follow the recommendations of the early research closely enough.

**New institutional environment due to standards and accountability movement.**

Another limitation in applying the early research on school size to the current urban small high school reform movement is the fact that the standards and accountability movement, with its emphasis on upgrading the curriculum, has fundamentally changed the institutional context in which schools and students are making curricular decisions. Contemporary secondary schools are experiencing vastly different institutional pressures—from high-stakes exit exams to pressures to get all students through a college preparatory sequence of courses—than their predecessors from the 1980s. Thus it is possible that the current institutional environment is influencing the relationship between school size and student course-taking in ways that were not evident when the early school size research was conducted.

Traditionally, institutional theorists have viewed educational organizations as “loosely-coupled” in the sense that the technical core of teaching and learning is buffered from institutional pressures by layers of bureaucracy (Bidwell, 1965; Weick, 1976; March & Olsen, 1976; Meyer & Rowan, 1978). Writing in the 1970s, these authors argued that because there was no reliable evaluative tool to assess how much schools were actually teaching students, the legitimacy of schooling depended on maintaining public confidence without any actual proof of learning. Thus the field of education has relied on administrative mechanisms like certification, accreditation, and bureaucracy to maintain legitimacy. Organizational structures were based on rule-like assumptions about how to organize—“rationalized myths”—rather than on technical efficiency (Parsons, 1960;
Meyer & Rowan, 1978; Dimaggio & Powell, 1983). This was because educational organizations of that era were operating in a “weak technical, strong institutional” environment (Meyer & Scott, 1983). But concerted institution-building by education professionals, government agencies, and private sector organizations over the intervening decades has begun to produce a more elaborate technical environment for schooling, an increasingly sophisticated theory of education productivity, and a technical capacity to inspect instructional outcomes (Rowan & Miskel, 1999; Fusarelli, 2002). Thus, schools now experience much stronger demands for technical performance without a decline in demands for institutional conformity (Rowan & Miskel, 1999).

Today, American secondary schools experience more institutional pressures than at any other point in our history. These pressures take the form of federal, state, and local mandates, and testing regimes, which (albeit imperfectly) convey to the public a measure of how well schools are educating their students. High-stakes education policies such as graduation requirements and high school exit exams operate by generating institutional pressure for high schools to ensure that their students take the courses that are needed for them to successfully complete the exit requirements. These policies have clear implications for the structure of the curriculum. By forcing students to pass tests associated with core courses in order to graduate, they strongly encourage schools to eliminate the “bottom track” of lower-level courses (Bishop & Mane, 1998). Interestingly, the fact that all schools, regardless of size, are feeling these same pressures could theoretically reduce the positive course-taking effects we might expect to see in small schools due to the “constrained curriculum” hypothesis. The standards movement was supposed to reduce between-school variation in course-taking in the same way that small school reform was supposed to reduce within-school variation in course-taking. In an era where a college-preparatory curriculum for all students is the new gold standard, and all schools are under direct pressure to get their students through core academic courses, the indirect curricular advantages of small school size may be washed out.
Research conducted on school size during an era when the standards and accountability movement was still nascent, such as the early research from the 1980s and 90s, may have limited applicability to contemporary small high school reform simply because the increased institutional pressures on schools have altered the relationship between schools’ size and curricular structure. I will discuss the current New York City institutional context in Chapter 3.

In sum, I have highlighted four reasons why the early positive research from the 1980s and 90s, which linked small school size to improved academic outcomes, may not be generalizable to contemporary small high school reform. I now turn to the more recent research, conducted since 2000, which does directly evaluate the current reform.

Recent Research on Small Schools—Increased Validity but No Focus on the Curriculum

The concern that the early research on the link between school size and academic outcomes may have both overstated the relationship between school size and student outcomes and may have limited applicability to the contemporary, urban small school reform appears to be born out in the more recent research, which is directly evaluating the current era of small school reform. Indeed, now that sufficient time has passed since the first new urban small schools were created (ranging from the 1990s to the early 2000s, depending on the location), scholars have begun assessing the impact of the reforms themselves. In this section I will focus on small school research conducted from 2000 to the present. By the latter years in this period, scholars began to use more sophisticated analytic techniques to address the concern of non-comparable samples between small and large schools. First I will review the recent studies evaluating small school reform and discuss their implications, and then I will outline what we still don’t know.

At the turn of the 21st century researchers began to examine the academic results of small schools created through active reform policies. This is an important departure from the early
research that focused more broadly on the associations between school size and student outcomes. The first cadre of these studies, while an improvement over the national studies in that they focused on schools that were small by design, still suffered from methodological flaws limiting their ability to fully address the concern of selection bias and thus infer causal effects. These studies largely supported the positive findings of the early work. Research in both Chicago and New York City—both school districts that experimented with small school reform in the 1990s—found better student outcomes in smaller schools (Lee & Loeb, 2000; Wasley et al., 2000; Darling-Hammond, Ancess, & Ort, 2002).

A good example of this point is found in the Darling-Hammond, Ancess, and Ort (2002) study of students’ experiences in the new small schools that were designed to replace a failing comprehensive high school. The authors found that the new schools, which enrolled a more “at-risk” student body than the previous school or the city as a whole, had lower drop-out rates, higher college-going rates, higher reading and writing Regents test scores and comparable math Regents test scores and SAT scores when compared to students in state-designated “similar schools” (schools serving similar populations). Yet, large-scale national studies like the one reviewed in the previous section are not the only types of study that suffer from unmeasured bias—it can also cloud the results of studies such as this one, which focus on individual schools and school districts.

For example, if students in school districts where there are newly created small schools have an element of choice in which school they attend, which is often the case, then students who end up in the small schools may not be equivalent to those who do not (Jennings & Pallas, 2010). If the positive results associated with small size are actually caused by selection of students with positive unmeasured traits such as motivation, then downsizing all high schools would not achieve the same results. I now turn to recent studies on small school reform that address the selection bias concern and consequently support more believable causal inference. Tellingly, these studies have failed to
find as consistent positive academic outcomes associated with small schools as the older research on school size from the 1980s and 90s or the early studies on contemporary small school reform, such as Darling-Hammond et al. (2002), which do not adequately address selection bias.

The first study straddles the useful, but somewhat artificial delineation I have drawn between the early and recent research bodies, as it employs sophisticated analytic techniques, but with a national dataset (the Educational Longitudinal Study of 2002). The authors compared results from a HLM approach to those from a propensity score matching approach. They found little effect of attending a small school on math achievement with either analytic approach, however the HLM estimates were somewhat larger than the matching estimates for post-secondary expectations and number of colleges to which students applied. However, there are several limitations to the applicability of this study to urban small school reform. First, while the more recent time frame may allow for some small schools created through purposeful reform to be included in the national sample, it is still a study of school size writ large, not urban small school reform. Furthermore, for the propensity score matching approach, the authors’ identified small schools as those with enrollments of between 800 and 1,999 students (again, substantially larger than what is being advocated by current urban small school reformers).

As in New York City, Chicago has extensively implemented small high school reform, also supported in large part by the Bill and Melinda Gates Foundation. Between 2002 and 2007, the Chicago High School Redesign Initiative (CHSRI) created 23 autonomous new small high schools to replace larger, closed schools. Recent analysis using multilevel modeling, found mixed results: improvements in engagement-related indicators such as attendance and graduation rates, but no difference in average ACT scores as compared to similar students in similar schools (Sporte & de la Torre, 2010). In some years, the average GPA of students in CHSRI schools was slightly higher than
their counterparts in other Chicago schools, but it remained less than a C average and was not significantly different in the most recent year (2007-08).

The most methodologically rigorous evaluation of urban small schools to date was conducted by MDRC (Bloom et al., 2010). It focuses on academically non-selective “small schools of choice (SSCs)” in New York City, which enroll fewer than 550 students and were created after 2002. The authors took advantage of lotteries, instituted by the Department of Education’s high school application processing system starting with the class entering ninth grade in 2005, for some seats in oversubscribed schools. When they are in the eighth grade, public school students in NYC must rank in order of preference, up to 12 high schools they would like to attend (for more detail on high school choice process see Chapter 4). Each year, some schools have more applicants than seats available. When this occurs, a lottery is created that randomly determines which students are assigned to that school. Bloom et al. exploited this random assignment to estimate the causal effects of attending SSCs for four cohorts of students who entered high school in the fall of 2005, 2006, 2007, and 2008, respectively. The sample of schools included in this study consists of 105 SSCs, which were oversubscribed. Students not lotteried into any SSC who attended another NYC high school, not necessarily a large one, served as the control group.

Students in one cohort (2005) were followed for four years and had statistically significantly higher graduation rates (6.8% higher) if they enrolled in SSC’s than if they did not. Subsequent cohorts had more credits toward graduation each year in SSCs than in other schools and were less likely to fail a course in a core subject. Finally for the 2005 cohort, Bloom et al. examined SSC effects on the percentage of students who pass the Math A or English Regents examinations with scores of 75 points or higher (the cutoff for exempting incoming students at the City University of New York from remedial courses in these subjects). The authors found no observed effect in math and a 5.3 percentage point increase in English. Thus, the MDRC study supports the conclusion that
smaller schools improve graduation rates and provides some limited evidence on their effectiveness for college readiness in English. However while this study provides the cleanest estimates of small school effects to date, it is limited in its generalizability—it only evaluated new small schools that were oversubscribed for some of their seats. This is problematic because the schools that are oversubscribed are likely to be the higher performing of the new, small schools, meaning that the results from this study may overstate the positive impact of the reform.

A recent quasi-experimental study of the 22 new small schools in Chicago uses variation in the distance between students’ homes to high schools in an instrumental variable (IV) framework to evaluate the effect of attending a small high school on student performance (Barrow, Claessens & Schansenbach, 2010). In their IV results, the authors found a positive effect of small school attendance on continuation through high school and graduation, but not on academic achievement as measured by test scores in 9th and 10th grade or on the ACT. These findings are consistent with the pattern from recent research that new urban small schools are improving engagement-related outcomes but are having less impact on academic outcomes. However, by measuring achievement after only one or two years of high school, this study does not allow students to receive the full four-year “treatment” of attending a small school. None of the studies on small schools in Chicago have examined students’ curricular experiences or the relative rigor of courses in small versus large high schools.

Finally, Schwartz et al. (2011) utilized a similar approach to the Barrow et al. (2010) study (using the distance between the nearest small school or large school and a student’s home as an instrument) except in New York City. Using the same 550-student cutoff as the Bloom et al. study, the authors examine the impact of small schools from both the most recent (post 2002) and past waves of small school reform (fully created before 2002). Interestingly, in their IV results, they find that only the new (post 2002) small schools produced better graduation rates and passing rates on
the English or math Regents exams than larger high schools. The earlier wave of small schools actually produced worse results on both graduation rates and rates of passing the Regents exams. Sub-group analyses conducted by the authors found few differences and did not support the theory that small schools produce differentially better effects for disadvantaged students. These results suggest that size alone may not be the critical characteristic in improving student outcomes but leaves open to question what is different about the new small schools, if not their student composition or resources. My study will build on this research by assessing whether differences in curricular organization between the different waves of small school reform in New York City may lend insight into these differential results.

When drawing conclusions from the research on small school reform in Chicago and New York City, it is important to keep in mind that contemporary small high school reform in both cities was part of broader, citywide reform initiatives. In Chicago, the small schools generally emphasized development of a clear vision to guide work and measure progress, a student-focused and rigorous curriculum, data-driven decision-making, an emphasis on personalization, professional development, and communication with parents and stakeholders (Wasly et al., 2000; Darling-Hammond, Ross, & Milliken, 2006). And as I discuss below, during the recent wave of reform in New York City, small schools were explicitly designed to improve academic rigor, foster personalization, and utilize community partnerships (Hemphill & Nauer, 2009; Quint et al., 2010). They also received start-up support, increased autonomy over decisions like hiring, and temporary exemptions from accepting English Language Learners and Special Education students. Darling-Hammond et al. caution that while these design features are shared by many small urban schools and have been found to produce positive outcomes for low-income students of color, they are not necessarily shared by all small schools. As the Schwartz et, al study indicates, simply reducing school size, without putting in place these attendant features, may be less effective.
Summary of recent research. To sum up, the research focusing on schools created through recent urban small school reform initiatives has supported the link between school size and student engagement related outcomes such as attendance and graduation, but has failed to find convincing evidence that small schools facilitate greater student learning than the existing alternatives. Moreover, these studies affirm the concern that the findings from the early research on school size may not translate to an active small high school reform. Finally, the Schwartz et al. (2011) study calls attention to the danger of conflating school size effects with the effects of other features of the small school reform package in New York City such as extra funding, increased autonomy, or the start-up energy that can accompany the creation of a school, but which may fade as time progresses.

What We Still Don’t Know

Taken together, the recent research on small school reform has been informative, but has also left many crucial questions unanswered. We know that small schools are not the silver bullet policy reform that many had hoped. Improved student achievement, while evident in some schools, is far from a guaranteed outcome of reducing school size. What might account for these lack-luster academic results? Why do the newer small schools in New York City seem to be doing better than the older ones? Do students’ experiences in small schools depend on their initial achievement levels or socio-demographic characteristics? What is the relationship between the constrained curriculum that necessarily accompanies smaller schools and the opportunities that students have to learn within them?

In order to answer these types of questions we must examine not just the final outcomes of education (test scores, graduation rates) but also the engine of learning itself—the curriculum. The original hypothesis for the link between school size and student learning was that smaller schools
would be more likely to implement a constrained academic curriculum with few low-level courses. In other words, the idea was that the curriculum would be constrained \textit{upwards}, to a rigorous, college preparatory level of difficulty. If small high schools do indeed implement this type of curriculum, then all students, particularly those from disadvantaged backgrounds, would be more likely to take a college preparatory sequence of courses in small schools because the constrained curricular structure would give them few other options. But despite this underlying theory, no studies to date have evaluated the curricular structures or student course-taking patterns in the new small urban high schools. Thus, there is no solid evidence that when small schools constrain their curriculum it is necessarily constrained upwards in terms of difficulty.

My dissertation builds on the existing literature by interrogating the relationship between school size, curricular organization, and students’ math course-taking in both new and existing small schools in New York City. Furthermore, I will address whether small schools produce differential course-taking outcomes for students of different initial ability levels and socio-demographic backgrounds. In the next section I will situate small high school reform within the broader ideological debate surrounding the curriculum. In the existing literature, small school reform has largely been disassociated from the history of school size and the curriculum. By locating the current reform within the evolving relationship between school size, curricular structure, and equity throughout the history of schooling in America, I am able to apply a historical lens to understanding why the contemporary reform may not have achieved all of its goals. This understanding will also help inform my conceptual framework (discussed in Chapter 3) of the pathways through which school size might influence students’ experiences in the math curriculum.

**Historical Overview of the Link between Curriculum, School Size, and Equity**

The content and form of the curriculum has long been a contentious issue in the history of the American education system. Educators, parents, school officials, and researchers continually
grapple with big curricular questions such as what students should learn, and whether all students should learn the same skills and content or whether course work should be differentiated according to students’ ability, motivation, and future aspirations. These questions have been debated since the inception of secondary schooling in America (Kliebard, 1986). Average school size co-evolved with the changing curricular priorities throughout history but until recently was more of a practical consideration linked to student population size and facilities or as a means to achieving curricular goals (i.e., large schools have more capacity to provide a differentiated curriculum than small schools).

On one side of the curricular debate is the view that all students—regardless of their educational abilities or aspirations—should experience an intellectually challenging academic curriculum that prepares them equally well for college or work. This view, first codified in the 1893 Committee of Ten report (discussed in more detail below), stems from the assumption that all students’ academic needs are similar and therefore they should be provided with a constrained academic curriculum that does not differentiate students by social background, ability, or future plans (National Education Association, 1893). On the other side is the view, first espoused in a 1918 report entitled the Cardinal Principles, that schools are responsible for sorting students according to their intellectual ability and motivation and matching them to their future social and economic positions (National Education Association, 1918). From this “social efficiency” perspective, secondary schools should offer a wide range of academic and vocational courses that vary both in content and rigor (Labaree, 1999; Lee & Ready, 2009).

Interestingly, our national response to the question of whether all students should learn the same thing (or at least our rhetoric on this issue) has been reminiscent of a pendulum, swinging back and forth between these two ideological stances. And because school size determines, to a large extent, a school’s capacity to efficiently offer a differentiated curriculum, our concept of the ideal
school size has swung with it. But what is also evident in the historical record is that students’ socio-demographic characteristics have remained crucial factors in their access to high-status knowledge throughout curricular paradigm shifts and the attendant fluctuations in average school size. Inequalities in access to high-status curriculum have historically manifested themselves in the guise of race, class, and achievement-based stratification both within-schools in the form of curricular differentiation and between-schools in the form of unequal academic intensity both in terms of which courses are offered and the rigor of the courses that are available.

The persistence of curricular stratification and inequality in access to high-status knowledge throughout the history of American education has important implications for contemporary small high school reform, which seeks to reduce these inequities by improving disadvantaged students’ curricular experiences. This study will be the first to situate small school reform within the broader history of curricular stratification in the United States, giving it a unique perspective into the challenges facing the reform. In this section I will review the history of conflicting curricular viewpoints and their connections to school size while emphasizing the importance of students’ socio-demographic characteristics and schools’ student body composition in determining curricular opportunities. Then I will briefly review the research on curriculum effects and delineate the potential sources of curricular inequality.

A Legacy of Curricular Stratification

The history of US public education has been a constant struggle over what should be common in the education of all youth and what should not (Hammack, 2004). In some eras we have placed precedence on schools being common in the sense that all students in a community should attend the same school regardless of their race, class, or future aspirations. In other eras we have emphasized that the curriculum should be common so that all students, regardless of what school they
attend, experience the same academically oriented course of study. In struggling with these fundamental concerns, education reformers throughout the ages have tinkered with the size and internal organization of schools to accomplish one or the other of these goals. In other words, increases or reductions in school size have historically been used as means to accomplish curricular or demographic educational ends, rather than ends in themselves.

And while the goals of common schools and a standardized curriculum are not mutually exclusive, as schools became more inclusive of students of different backgrounds, they concurrently became more internally stratified through increased curricular differentiation (Oakes, 1985; Lucas, 2001). One interpretation for this phenomenon is the theory of Effectively Maintained Inequality (EMI), which posits that socioeconomically advantaged individuals secure for themselves or their children some degree of advantage whenever possible. If quantitative advantages are common, the socioeconomically advantaged will obtain quantitative advantage; however in the absence of quantitative distinctions, the socioeconomically advantaged will obtain qualitative advantage (Lucas, 2001). When applied to educational attainment, EMI holds that so long as a particular level of schooling is not universal the upper class will employ their advantage to secure that level of schooling. However once that level of schooling becomes nearly universal, the socioeconomically advantaged will seek out qualitative advantages at that level to secure better education (and thus a competitive advantage) for their children (Lucas, 2001). EMI provides a useful lens for understanding the historical relationship between the curriculum and equity of educational opportunity. It is also helpful for illuminating the present-day challenges that small-school reform faces in attempting to equalize access to a rigorous curriculum and improve college-readiness among disadvantaged students.

**Common schools, common curriculum—for a select group.** The first major national report on the high school was issued in 1893 by the Committee of Ten, a group of college presidents
and professors, who sought to increase the standardization of preparation for higher education (National Education Association, 1893; Tyack & Cuban, 1995; Kliebard, 1995). The authors argued that students should be permitted little curricular choice and that all high schools should offer a narrow academic curriculum that did not differentiate between “college” and “non-college” students, but rather provided a rigorous academic training to all (National Education Association, 1893). This position on the role of schooling in America was reflective of the dominant educational ideals at the close of the 19th century. Schools were viewed as essential to nation-building, training a productive and civic-minded citizenry, and reducing social inequalities (Cremin, 1980).

But rhetoric aside, access to secondary education in this era was severely limited—only one in ten adolescents between the ages of 14 and 17 were enrolled in high school in 1900 (Tyack & Cuban, 1995). At the turn of the 20th century, high schools served a small, socio-economically advantaged, largely white, and disproportionally urban portion of the population. From a social theory perspective, while the ethos of democratic equity is commendable, high schools in this era were not needed to perform the function of sorting and allocating students to their future roles in society because only an already select group of children entered their doors to begin with (Sorokin, 1927). In other words, inequality in access to high-status knowledge and the competitive edge that it confers was already guaranteed by the fact that only the socioeconomically advantaged attended high school during this era—qualitative distinctions such as differentiated courses of study were not yet needed (Lucas, 2001).

Finally, high schools of this era were also quite small in size; one third of high school students in America had only one to three teachers and two thirds had one to ten teachers (Tyack & Cuban, 1995). Small staffs such as these did not have the capacity to offer a heavily differentiated curriculum. Put another way, the small schools of this era were generally small by default rather than by design and within-school curricular uniformity was often as much a necessity as a conscious
choice. So while the high schools of this era were indeed relatively “common” with respect to both aspects of education—demographics and curriculum—for those students who attended high school during that era, they were not inclusive of the majority of high-school aged youth, particularly those who were economically disadvantaged, non-white, female, or from rural areas.

**Diversity and differentiation.** In the decades after the Committee’s report, the American economy and its school system were transformed by the massive forces of immigration, urbanization, and industrialization, which would help bring the large, comprehensive high school model to dominance (Powell, Farrar, & Cohen, 1985). A new vision for the high school curriculum rose to prominence, codified in a report entitled *Cardinal Principles of Education* (National Education Association, 1918). The authors argued that high school students’ coursework should be tailored to their future occupational and educational plans. The underlying rationale was that of “social efficiency”—schools should prepare students for their future adult roles (Labaree, 1999). The resulting educational recommendations encouraged secondary schools to offer a wide range of academic and vocational courses that varied both in content and rigor. The reformers argued that providing only traditional academic curricula was unfair to students because it ignored the reality that students enter high school with different academic skills and aspire to different occupational futures (Kliebard, 1995; Lee & Ready, 2009).

Tellingly, the social efficiency view of secondary education first rose to prominence in an era when the American high school was experiencing a rapid growth in enrollment, fuelled by both an influx of new immigrants and a surge in enrollment from sections of the American populace that had previously never attended high school (Cremin, 1980). A new curriculum, the argument went, was needed to engage a new type of student—students of social and ethnic backgrounds that previously had never attended secondary school. The assumption was that the new students entering high schools were either uninterested or incapable of rigorous academic coursework and that
differentiation was the key to universalizing high school enrollment and keeping the new wave of academically uncommitted students in school through graduation (Tyack & Cuban, 1995). This is in keeping with the concept of Effectively Maintained Inequality; once high school enrollment became nearly universal, the socioeconomically advantaged parents were forced to seek out qualitative advantages such as “academic” or “college-preparatory” track placement for their children in order to maintain a competitive advantage (Lucas, 2001). The social efficiency argument also manifested itself in the physical organization of schools. Small urban schools were inefficient from a financial perspective and could not support the wide range of course offerings that the early twentieth century reformers desired so new, larger schools in which more students could be educated for less money were modeled after contemporary factory models of low-cost assembly line production (Wells & Holme, 2005).

Writing in the 1950s—40 years after the Cardinal Principles—James B. Conant (1959) continued to advocate for the comprehensive high school, arguing that it was democratic because it housed students of all backgrounds and future aspirations under one roof and that the key to its continued success was to strengthen curricular differentiation. Labarce (1999) argues that Conant’s view of the ideal American high school—inclusive, yet internally inequitable—was popular and enduring because it was the product of both politics and markets: it offered universal access and differential individual payoff. The tension between these two conflicting pressures was internally contained through a highly stratified curriculum tracking system.

Thus, during the first half of the 20th century schools in the North were increasingly “common” in the sense that students of diverse backgrounds generally attended the same schools (although some between school segregation by race and class certainly remained), however they were very uncommon in the sense that students were segregated within schools such that students of different background rarely shared the same classrooms and were rarely taught the same curricula. In
the South, schools were neither common in terms of demographics due to rigid racial segregation policies, nor with regards to the curriculum due to unequal resources for black schools.

**Civil Rights Era—an era of contradiction.** In many ways the 1960s and 70s were a time of triumph for racial/ethnic minority and feminist groups. In particular, civil rights activists won stunning victories in the courts that began to dismantle the de jure racial segregation policies of the South. But just as racially desegregated schools were becoming a reality in the South, they were becoming more and more out of reach in the North. Post-World War II suburbanization had transformed America's metropolitan areas into segregated doughnuts, with most blacks in the urban center and whites increasingly in a suburban ring (Massey & Denton, 1993; Orfield, 1996). Furthermore, even when schools were integrated by race and class, they rarely ensured that students of different backgrounds actually experienced schooling together (Eyler, Cook & Ward, 1983; Lucas, 1999). In fact, case studies document that the typical bureaucratic response to increased student diversity accompanying integration was to create programs that re-segregated students within schools, usually by expanding the number of non-demanding curricular offerings (Grant, 1988; Cusick, 1983).

Also beginning in the 1960s, the rigidity of the curricular tracks began to break down (Lee & Ready, 2009). Whereas high schools had previously determined students’ track placement early and permitted minimal movement between academic, general, and vocational programs, 60s and 70s era high schools increasingly allowed students to choose their own courses. But while tracking was less deterministic, high schools remained large in size and curriculum differentiation remained high. Students were generally able to choose how far they would advance through the subject “pipelines” as well as the difficulty of their courses (i.e. general, honors, AP).

However, empirical research in comprehensive high schools indicates that the choice-driven curricular model often results in a sort of “de-facto” tracking of students which reproduced the
same socio-demographic inequalities associated with traditional tracking (Powell, Farrar, and Cohen, 1985; Yonezawa, Wells & Serna, 2002; Lucas & Berends, 2002). These studies found that within-school stratification was partially due to the fact that students from more socio-economically advantaged backgrounds (often guided by savvy and sophisticated parents) chose more demanding courses, while students from less advantaged and immigrant families (often lacking this type of parental involvement) would either choose, or allow themselves to be guided toward low-level academic courses.

In short, the 1960s and 70s were decades in which the nation struggled over the issue of who should go to school with whom, but even in the locations where students of diverse backgrounds attended the same schools, they rarely were placed in the same classrooms or had access to a common curriculum due to within-school stratification which was facilitated by the differentiated or “comprehensive” curriculum. Furthermore, in the public school systems of many metropolitan areas, whole schools and school districts were becoming segregated by race and class as white suburbanization intensified, beginning a trend of increasing between-school stratification that would continue to gain steam in subsequent decades.

The standards and accountability movement—a return to the common curriculum? Beginning in the late 1970s and early 1980s and aided by a few high profile national articles and reports, the pendulum of national opinion began to swing back towards the view that all students should follow a prescribed, academic course of study—a common and rigorous curriculum became the focus of secondary school reform, rather than a common school (Adler, 1982; NCEE, 1983; Boyer, 1983; Goodland, 1984; Sizer, 1984; Powell, Farrar, & Cohen, 1985; Oakes, 1985; Hampel, 1986; Gamoran & Berends, 1987). The most influential of these reports was *A Nation at Risk*, produced by the U.S Department of Education. Apocalyptic in tone, this report pinpointed high schools as the weakest link in a wholly inadequate education system and declared the differentiated
high school curriculum to be enemy number one. “Secondary school curricula have been homogenized,” the report proclaimed, “and diffused to the point that they no longer have a central purpose. In effect, we have a cafeteria-style curriculum in which the appetizers and desserts can easily be mistaken for the main courses” (p.18).

The report urged states and local districts to mandate graduation requirements such that all students must take 4 years of English, 3 years of mathematics, 3 years of science, 3 years of social studies, and half a year of computer science. The NCEE believed that all students should follow the same, academically oriented, course program regardless of their social backgrounds, academic ability, or future plans. For the commission, equality of educational opportunity no longer meant treating different students differently, but rather meant that all students should have access to the same high-quality academic courses (Angus & Mirel, 1999).

Thus, even as desegregation was increasingly becoming a politically untouchable, “third-rail” policy strategy for achieving educational equity, constraining and “up-grading” the high school curriculum became a hugely popular cause, drawing support from a broad and diverse constituency (Angus & Mirel, 1999). The most popular policy instrument used to achieve academic upgrading was raising state graduation requirements (Clune & White, 1992). Three years after A Nation at Risk, 45 states had increased their high school graduation requirements, 42 states had increased their math requirements, and 34 had increased their science requirements (Stedman & Jordan, 1986). And while the new state graduation requirements generally did not have a substantial impact on the students already in college preparatory tracks (who were already taking the newly required courses), it did dramatically influence the course taking patterns of lower-track students. Enrollments in vocational education, for example, declined significantly during this period as students who had previously taken those courses were forced to increase their participation in core academic subjects. By the end
of the decade, roughly one quarter of students completed an extra year of math, and one third completed and extra year of science (Lee & Ready, 2009).

At a national level, support for constraining and upgrading the curriculum was exemplified by an unprecedented meeting in 1989 of President George Bush and the nation’s governors (led by then governor of Arkansas Bill Clinton) in which they drew up a list of goals geared towards revitalizing the American education system. Goal Three focused on improving competency in academics subjects including English, mathematics, science, history, and geography. Goal Four focused specifically on rising to first in the world in math and science achievement. This first bipartisan meeting set into motion what would come to be called the “standards and accountability” movement, which would be enacted both at the state level, through high-stakes graduation requirements and exit exams, and at the national level through the federal No Child Left Behind (NCLB) Act of 2001 and the more recent Race to the Top program of 2009—a $4.35 billion dollar initiative to raise student performance by stimulating innovation and infusing accountability into the nation’s public school systems.

Central to the standards and accountability movement, or at least the rhetoric surrounding it, is the view that all students—regardless of their social background, where they attend school, who they attend school with, and what size their school is—should have access to a rigorous, college preparatory curriculum and that “achievement gaps” between high-income and low-income and white/Asian and black/Hispanic students should be closed. One of the major goals both of Race to the Top, and the accountability movement more broadly, has been identifying and implementing strategies to turn around the country’s lowest-performing schools.

But despite these goals, our national commitment to ensuring equity and curricular rigor has only extended to policy reforms that address the problem within the current hyper-segregated school context. Racial desegregation has both lost public support and has been abandoned by the
courts (*Parents Involved in Community Schools v. Seattle School District No. 1*, 2007), desegregation by social class or achievement-based indicators has gained little traction (Kahlenberg, 2007), and inter-district transfer programs that would allow disadvantaged urban students to attend well-resourced, academically rigorous school in the suburbs are few and far between (Wells, Warner, & Grzesikowski, 2013). Thus, due to both legal and political constraints, education policy makers have sought new strategies to decrease educational inequities between advantaged and disadvantaged students that do not require them to attend the same schools. Urban small school reform has flourished. By the early 2000s small schools and schools-within-schools had been endorsed by high-profile developers of whole-school reform models such as the Coalition of Essential Schools, the Center for Research on the Education of Students Placed at Risk, and the Institute for Research and Reform in Education (Quint et al., 2010). Moreover, beginning in the 1990s and continuing through the first decade of the 21st century, the philanthropic community has strongly supported small schools; the Annenberg foundation, the Carnegie Corporation, and the Bill & Melinda Gates foundation have all invested heavily in the creation of small high schools. Finally, the federal government has supported small school reform and even has a “Smaller Learning Communities Program” which awards funds on a competitive basis to large high schools who want to reorganize into smaller “schools-within-schools” or themed “houses” (Lee & Ready, 2007).

The result of this current policy paradigm is that strategies such as small school reform that seek to aid disadvantaged students must do so within the current context of intense pressure to meet proficiency standards based on student test scores combined with between-school (and often between-district) stratification by prior academic achievement, race, and class. In many cases, small schools are being introduced as replacements for large schools that were characterized by concentrated poverty, weak instructional capacity, and chronic low performance. This is crucial because if the relationship between school size and curricular offerings is mediated by student body
composition, then between-school stratification may influence the academic results of small school reform. Furthermore, any policy reforms attempting to equalize the academic intensity of students’ curricular experiences, must not only reduce within-school inequities but must also bring the curricula of segregated, high-poverty schools up to the standards of schools of concentrated privilege.

I now turn to a brief review of the history of small school reform in New York City. Situating New York City small school reform within the broader historical overview of the curriculum, school size, and educational equity is helpful for understanding the goals and assumptions of the city’s distinct waves of small school reform. In some ways the New York City experience with small schools is a microcosm of the national experience. However, New York City’s experimentation with small schools movements has historically been, and continues to be, more extensive in scope than anywhere else in the nation, making it an excellent location to evaluate the potential of this reform. Moreover, even though New York City is unique, much of what I will learn from studying small high school reform in the school system that has implemented it most aggressively will translate to urban small high school reform across contexts. Finally, reviewing the history of small schools in New York is also helpful from a practical perspective, as it provides concrete information that I will need to construct an appropriate methodological approach for addressing my research questions.

**Three Waves of Small School Reform in New York City**

In many ways, New York City is the epicenter of the small schools movement in the United States. While small schools and other small learning environments have been instituted in urban school districts across the nation, none have surpassed the scale and rapidity of adoption that has taken place in New York City since 2000. The vigor with which the small school reform has been implemented in NYC is due largely to the shared commitment of Mayor Bloomberg, the NYC
Department of Education and the Bill and Melinda Gates Foundation, which has invested more than $150 million to support the development of small schools in NYC alone. However, the small schools movement in New York City did not begin in 2000 with Bloomberg, Klein, and the Gates Foundation. Indeed, New York City is also considered by some to be the “birthplace” of the small schools movement (Jennings & Pallas, 2010) and has hosted three separate waves of small school creation over the course of the past 50 years.

**Small “Alternative” High Schools—1960s and 70s.** Beginning in the late-1960s and early-1970s in New York City and the nation, dissatisfaction with the large comprehensive high schools which predominated during this era, in which disadvantaged students often found themselves at the bottom of a highly stratified curriculum, fueled an alternative or experimental schools movement geared towards helping students who were not succeeding in the traditional school system and were in danger of dropping out. Centered in New York City, many of these schools were small in size and specifically designed to foster a sense of community (Hemphill & Nauer, 2009). The alternative schools were often countercultural in that students as well as staff had a hand in developing the curriculum and courses tended to have a left-leaning political slant (Hemphill & Nauer, 2009). New York’s Central Park East Secondary School, founded by Deborah Meier in 1984, exemplifies the type of small schools created during this era. Students at Central Park East “might study the Vietnam War from the point of view of the Vietnamese or the black power movement or the history of women factory workers in World War II” (Hemphill & Nauer, 2009).

But while many of the goals of the alternative small schools are the same as the contemporary small school reform—a more intimate environment, better student engagement, fewer drop outs, increased relevance to students’ lives—the curricular goals of the two movements are somewhat divergent. The alternative schools were focused on co-creating a new more culturally-relevant curriculum and often rejected the notion of standardization. These schools were granted
increased flexibility with regard to staffing, the structure of the school day, assessment, and resource allocation (Jennings & Pallas, 2010). They were even permitted exemption from standardized testing. So while the founders of the alternative small schools were similar to contemporary reformers in that they targeted disadvantaged and minority students, their underlying assumptions and ultimate educational goals were very different, arising from a more sociological rather than outcome-driven view of what small learning environments could achieve. According to a recent report by Hemphill and Nauer (2009), several of these original small schools are still flourishing, while others have either closed or are floundering. I will refer to small schools created during this era as “small alternative schools.”

**Small School Creation in the 1990s.** In 1992, then chancellor Joseph Fernandez launched a second, more formalized round of small school creation, announcing that the city would create 30 small high schools in partnership with intermediary organizations. The Aaron Diamond Foundation financed the initial start-up costs through a fund that eventually came to be known as “New Visions for Public Schools” (Jennings & Pallas, 2010). Schools were selected through a Request for Proposals process. The Fernandez era of small school creation also represented the first time that small schools replaced an existing, large high school (Jennings & Pallas, 2010). In the fall of 1992, Julia Richman High School stopped taking incoming 9th graders and the Coalition Campus Schools Project replaced it with six new small schools. Improved student outcomes at the “Julia Richman Educational Complex” were used to fan the fire for replacing more large schools, and the Campus Coalition Schools Project expanded to Monroe High School in the Bronx the next year (Bradley 1993; Jennings & Pallas, 2010).

The 1990s small school effort received a substantial boost in 1994 when the Annenberg Foundation invested $25 million in opening approximately 40 new small high schools through a coalition of four intermediary organizations, including New Visions, which came to be known as the
New York Networks for School Renewal. The new schools stressed small size, choice, autonomy, personalization, and the formation of professional learning communities of teachers and staff across schools (Quint et al., 2010). The small high schools of this era largely had selective admissions policies—not uncommon in NYC, but different from the large, comprehensive high schools in the city which are generally zoned. Nonetheless, they still served a larger percent of black and Hispanic students than the city as a whole (Quint et al., 2010).

In the mid-1990s, however, small school reform stalled as the chancellorship passed to Rudy Crew, who was unenthusiastic about the prospects of small schools. Crew disbanded the alternative high school division, the entity that administered the small schools, and increased enrollments at a number of the schools. By the time that Bloomberg took over as mayor in 2002, small schools enrolled fewer than 10 percent of the city’s high school students (Hemphill & Nauer, 2009) and nearly 70 percent of students were educated in high schools with enrollments greater than 1,400 students.

**Bloomberg Era Small School Creation—2000s.** During the Bloomberg administration there have been two distinct small schools initiatives: New Century High Schools (NCHS), which was launched in 2001, and the Children First initiative, which followed in 2003. NCHS was funded collectively by the Gates Foundation, the Carnegie Corporation, and the Open Society Institute to the tune of $70 million. Under NCHS, groups of educators submitted proposals to start new schools and were selected through a competitive process. Each school was supposed to have a thematic focus (technology, business, law, hospitality, etc.) and was required to partner with a community organization. New Visions for Public Schools served as an intermediary organization between the philanthropic organizations and the new schools.

In many ways, the small schools element of Children First can be viewed as a more aggressive (both in terms of speed and scope) continuation of NCHS’ small schools initiative.
However, the Children First policy was larger than just small school reform, it also involved systematic changes to the process of high school choice in New York City. With Joel Klein as the new chancellor of schools, the DOE moved aggressively to close large, underperforming high schools and replace them with new small schools (fewer than 550 students). In the 2002-03 school year there were 58 small high schools in NYC remaining from previous rounds of small school creation; six years later there were a total of 161 small high schools (Quint et al., 2010). I refer to the schools remaining from the early waves of small school reform as “veteran” small schools and the schools created during or after the 2002-03 school year as “new” small schools. Similar to the NCHS initiative, New Visions played the leading role in school creation under Children First. However, New Visions lacked the capacity to establish the number of schools needed for the initiative so the DOE and the Gates Foundation agreed that they would contract a total of 18 intermediary organizations.

The DOE decided to close high schools that graduated less than 45% of their students—generally, large, zoned, “comprehensive” high schools, which were concentrated in poor, non-white areas of Brooklyn and the Bronx. Between 2002 and 2010, 27 large high schools have been designated for closure (Jennings & Pallas, 2010). Once designated, these schools are then “phased out” in the sense that no new classes of freshmen are allowed to enroll but the remaining cohorts of sophomores, juniors, and seniors stay until they graduate. The logic behind the phase-out strategy was that it would allow the new small schools to grow gradually, adding a grade each year, and the remaining large schools would not be inundated with new students from schools that had been closed. The new small schools were made open to students, regardless of ability (“non-selective”) so that they could serve the same student populations that had previously attended the zoned, dysfunctional schools that were shuttered (Quint et al., 2010). In many cases, the new small schools were established at the sites of the large high schools that had closed or were in the process of being
phased out. In DOE jargon, these large school buildings, which now house multiple small schools, are called “educational campuses.” On average, four new schools were located on each of the 21 educational campuses where large or mid-size schools were closed (Quint et al., 2010). Amazingly, in 1992 there were only 99 high schools in New York City, by 2008 there were almost 400. Of the 300 or so newly created high schools, 75 were established during the first wave of the contemporary small schools (1993-2002) and over 200 were created during the second wave (2002-2008).

Summary

In this study, I examine the relationship between urban small school reform, the structure of the curriculum, and students’ course-taking patterns. In doing so, I have situated my analyses within an understanding of both the historical evolution of the link between school size and the curriculum in the United States as well as the existing literature on school size (and the shortcomings of applying it to the current context). By embedding the early academic literature on school size and the current small school reform in their respective historical contexts it becomes evident that findings from one might not generalize to the other. The awareness of this potential disconnect, combined with the lack of consistent positive academic findings from studies evaluating contemporary urban small school reform, leads me to question whether the “constrained curriculum” hypothesis generated by the early academic research on school size holds true for the current reform. I argue that a new, more complex theoretical model is needed, that not only conceptualizes the direct pathways through which school size might impact student course-taking but also includes student and school factors that might mediate this relationship. In the next chapter, I will discuss my conceptual framework for understanding the link between small school reform and students’ experiences in the math curriculum.
Chapter Three

Conceptual Framework

Both the academic literature on school size and the curriculum and the historical record on the relationship between these two features of schooling point to a positive association between high school size and the number and differentiation of course offerings. Furthermore, organizational theory posits that the internal structures of schools will become more complex and differentiated as they increase in size (Weber, 1947; Haller et al., 1990; Monk & Haller, 1993) and economic theory agrees that small schools cannot cost-effectively offer a wide range of courses and programs (Buzacott, 1982; Kenny, 1982). Thus, the link between small schools and a relatively narrow curriculum is well supported by both research and theory. However, a narrow range of courses constitutes only half of the constrained curriculum hypothesis. In order for small schools to both increase within-school equity and raise the overall academic intensity of their students’ coursework, they must constrain their curriculum upwards by offering a narrow curriculum consisting of rigorous, academic courses. Put another way, in the highly stratified context of today’s urban school districts, the only way that small schools can reduce inequities in disadvantaged students’ access to a college preparatory curriculum is to reduce between-school curricular disparities by offering their students a more competitive, advanced curriculum than they would have experienced in the large, closed schools.

I begin this chapter with a brief review of the research on curriculum effects and a discussion of the sources of curricular inequality. Next, I outline my conceptual framework for understanding the link between school size and student course-taking. I draw on my review of both the research on school size and the curriculum as well as the history of curricular stratification in the United States to argue that the assumption that small high schools will necessarily offer a
constrained and rigorous academic curriculum is overly simplistic, and fails to account for several important mediating factors such as student body composition, teacher quality, and the institutional context in which schools are operating, all of which can influence the relationship between small school reform and students’ curricular experiences. I will also discuss the possibility that the academic rigor of courses with the same title might vary across schools and outline how my study will address this concern.

A Brief Review of Research on Curriculum Effects

As discussed in the historical overview, most of secondary school education reform over the past three decades has focused on equalizing and upgrading high school students’ curricular experiences. The critical assumption behind these reforms is that students learn more in advanced courses and that the knowledge they acquire in these courses will help them to achieve better post-secondary outcomes (Attewell & Domina, 2008). Advocates of upgrading the curriculum also argue that inequality in access to academically intense courses in high school underlies much of current racial and socio-economic inequality in college attendance and perseverance (Adelman, 1999, 2006; Bowen, Kurzweil, & Tobin, 2005) The constrained curriculum hypothesis linking small school size to improved academic outcomes for disadvantaged students is also predicated on this assumption. A substantial body of research supports the association between a more advanced course of study and improved academic outcomes, as measured both by test scores and post-secondary success, even after controlling for students’ social and academic backgrounds (Alexander & Pallas, 1984; Adelman, 1999, 2006; Lee & Bryk, 1988, 1989; Lee et al., 1998; Gamoran & Hannigan, 2001; Natirello, Pallas, & Alexander, 1989; Horn, Kojaku, & Carroll, 2001). These findings are supported by related work linking curricular differentiation to stratification in students’ educational outcomes (Jones et al., 1986; Garet & DeLaney, 1988; Lee et al., 1988), and by the work linking the “constrained” academic
curriculum to more equitable and improved student learning (Bryk, Lee & Holland; Lee & Bryk, 1988, 1989; Lee & Smith, 1995).

However, it is difficult to draw causal inference in the case of curriculum research, because it is plagued by selection bias concerns at both the student and school level (Girotto and Peterson, 1999; Attewell & Domina, 2008; Lee & Ready, 2009). For example, students who choose and are accepted into more advanced courses are likely to be more academically skilled, prepared, and motivated, on average, than their peers who take lower-level courses. If this is true, then it is difficult to say with certainty whether students learned more and had better outcomes because they took more advanced courses or because they were the types of students that would exhibit these positive outcomes regardless of which course sequence they followed. Furthermore, schools that serve predominantly high-SES families are more likely to offer advanced course-work than schools serving a more disadvantaged clientele. As such, the correlation between taking advanced courses and improved student outcomes could be due to some other school factor such as student body composition, resources, or teacher quality.

Finally, some scholars argue that there is indeed a causal link between taking advanced courses and post-secondary success, but that it does not operate through increased student learning or content mastery but rather because advanced courses, which only enroll a small number of academically elite students, function as a signal to colleges of students’ merit and relative ranking (Collins, 1979; Labaree, 1999; Dougherty, 2006a; Attewell & Domina, 2008). This “Credentialist” thesis hypothesizes that, while taking advanced courses and post-secondary outcomes may be related at any given cross-sectional time point, as more and more students enroll in a specific advanced course, it will no longer serve as a distinguisher and the advantage gained from taking it will be diminished—the academic elite will be driven to find even more advanced courses through which to differentiate themselves from the pack. However, a recent study investigating this possibility, using
propensity score matching to address selection bias, did not find evidence to support the Credentialist thesis and affirmed that students do benefit from taking a more demanding curriculum, although the size of the benefits found were smaller and less consistent than the earlier research suggested (Attewell & Domina, 2008).

Nonetheless, there is some concern that the positive results of the curriculum research may not generalize to active policy reforms geared towards upgrading the curriculum for all students. Recent research from Chicago suggests that these concerns may be warranted (Allensworth et al., 2009). Chicago mandated college preparatory course work for all students in all high schools beginning with the students entering in 1997. In the ninth grade students were required to take Algebra I and English I (or a higher course in the sequence). Remedial courses were eliminated in both subjects. Yet despite decades of research linking a college preparatory course sequence to positive academic outcomes in high school and college, the reform did not have the desired effects. In fact, a rigorous study recently found that as a result of the reform, grades slightly declined, test scores did not improve, and students were no more likely to enter college (Allensworth et al., 2009). The authors question whether the Chicago Public Schools had the instructional capacity to implement this universal mandate in a way that maintained the quality of the newly required college prep courses. They note that many teachers who had taught remedial courses were suddenly required to teach college prep instead. Also classes became more heterogeneous in terms of student ability level. Research on successful de-tracking efforts has emphasized the importance of instruction and teacher buy-in for success in mixed-ability high school classrooms (Boaler & Staples, 2008; Gamoran & Weinstein, 1998). Whatever the cause, it is important to recognize that results from observational studies do not necessarily transfer to an active reform policy, particularly one implemented in a challenging urban context.
So while the exact impact of taking advanced courses on students’ educational outcomes is unclear, my study sides with the majority of researchers, educators, and policy makers by taking as a given that it is beneficial for students to take advanced and rigorous courses and that the analytic skills and content students learn in advanced courses help prepare them for both college and careers. I also take as a given that racial and socio-economic inequality in students’ course-taking is ethically problematic and likely contributes to the unequal post-secondary outcomes for low-income and minority groups. The question that this study sets out to answer is whether small school reform can lead disadvantaged students to take and succeed in more advanced courses.

Sources of Curricular Inequality

The constrained curriculum hypothesis holds that small schools can improve disadvantaged students’ educational opportunities without harming the opportunities of advantaged students if they reduce within-school curricular differentiation and ensure that all students have access to the same college preparatory courses. But it is important to acknowledge that inequalities in students’ opportunity to learn high-status knowledge do not arise solely from within-school inequalities in access to curriculum.

Research suggests that curriculum disparities can emerge from three sources: schools, courses, and classes (Cogan, Schmidt & Wiley, 2001; Oakes, 1990). First, students’ course opportunities vary from school to school as school leaders make different decisions about what type of courses will be offered. Historically, schools that serve predominantly low-income and minority students have offered fewer high-track classes when compared to affluent schools serving white and Asian students (Oakes, 1990). Second, schools can create another source of variation in opportunity to learn within their own buildings when they decide to offer different types and levels of courses to students in the same grade. This internal differentiation varies the curriculum from course to course
within schools. In racially and economically diverse schools, minorities and low-income students have historically been underrepresented in high-track courses. This is the type of curricular inequality that will be targeted directly by reducing school size. Finally, given a specific type of course, the curriculum can vary from class to class—classes that are purportedly at the same track level can turn out to be quite different with regards to both rigor and the quality of their resources and instruction. This type of variation can occur both within and between schools as teachers make different instructional decisions about the topics that will be covered and how much time will be spent on each topic. However, researchers have documented patterns in class to class variation of opportunities such that an Algebra II course, for example, at a school serving predominantly disadvantaged students is likely to be less rigorous and taught by a less-qualified teacher than a class with the same title serving advantaged students (Oakes, 1990; Coley, Ekstrom, Gant, Villegas, Mitchell & Watts, 1992; Meier, Stuart & England, 1989). Efforts to equalize access to a rigorous, college preparatory curriculum must minimize inequalities stemming from all three sources in order to be successful. This is an important point given the increasing stratification of students by race, class, and academic achievement between schools and school districts. To the extent that stratification occurs between schools or districts, increasing the within-school equity in course-taking by reducing school size and constraining the curriculum may have a limited impact on the system-wide level of equality.

**Equity, Adjusted?** The fact that racial and socio-economic disparities in advanced course-taking exist is undisputed. However, an important question is whether these disparities still exist after adjusting for students’ initial academic achievement. In other words, are racial and economic inequities in advanced course-taking the product of educator bias, or are they actually reflective of accurate assessments of students’ skills? Much of the research on teacher expectations concludes that, on average, teachers do have lower expectations for their minority and low-SES students as
compared to their white and affluent students, however these lower expectations reflect accurate perceptions of students’ academic performance (Ferguson, 1998, 2003; Attewell & Domina, 2008). Specifically with regards to course-taking, Lucas (1999) found that minority and low-SES students were disproportionately enrolled in lower-level coursework. However, once prior achievement was controlled for, he found that Black students were more likely than White students to take college prep math and English, and Hispanic students did not differ significantly from White students. However, SES differences in course-taking remained, even after controlling for prior achievement.

Disparities in academic achievement accrue over years. Can high schools be faulted for assigning students to different levels of courses based partially on their prior skills and content mastery? The answer is not clear-cut. For example, as I will discuss below, if some students enter high school having already taken and passed Algebra and some have not it seems only reasonable to assign the former group to Geometry and the latter to Algebra. Furthermore, requiring students with very weak academic skills to take a college preparatory curriculum without also implementing sufficient support mechanisms could lead more students to fail or drop out (Darling-Hammond, 2004).

On the other hand, research has documented the poor educational environments of low-track classes and does not support assigning initially low-achieving students to extensive low-level or remedial courses that are not geared towards helping them progress to more advanced, academic coursework (Gamoran, Nystrand, Berends, & LePore, 1995; Lucas, 1999; Oakes, 1985; Powell et al., 1985). Indeed, researchers and policymakers are increasingly suggesting a compromise, which would eliminate low-level courses while providing struggling students with a “double dose” of core subjects such as math and English to provide them with the necessary academic support to succeed (Nomi & Allensworth, 2009; Chait, Muller, Goldware, & Housman, 2007).
A starting point in understanding the role that small high school reform plays in shaping students’ curricular experiences is recognizing the enormous initial differences in students’ academic preparation upon entering high school. However, as I argue later, placing students into low-level courses based on their prior achievement, without putting supports in place to help them catch up is problematic and ensures that these students remain behind. For the purposes of this study, it is informative to provide both unadjusted “descriptive” information on schools’ course offerings as well as adjusted analytic results comparing the outcomes of similar students in small and large schools. The former provides information on students’ educational opportunities in terms of “access” to advanced coursework, while the latter evaluates how effective small schools are at getting their students through the math curriculum, given where they started. In the next section I will delve deeper into how schools’ curriculum structures interact with students’ social and academic characteristics to determine student course-taking patterns.

**Conceptualizing Student Course-taking**

High school math courses are typically organized hierarchically into sequences of increasing difficulty (Adelman, 1999; Schneider et al., 1998; Stevenson et al., 1994; Lee, Croninger, & Smith, 1997). For example, the lower end of the math sequence is composed of general math and Pre-Algebra classes and the higher end is composed of advanced courses such as Trigonometry and Calculus (Adelman, 1999; Schiller & Hunt, 2003; Schneider et al., 1998). Skills and concepts build on one another throughout the sequences, meaning that curricula in one class must generally be mastered before students can successfully move to the next course in the sequence. Course sequences in math can thus be conceptualized as sequences of educational opportunity, where a student’s opportunity to learn advanced content (Calculus or Physics, for example) is partially dependent on their successful completion of the preceding courses in the sequence.
Crucial to my conceptual framework is the well-grounded theoretical assumption that the two primary and interacting factors that determine how far students progress through curricular sequences or “pipelines” are; first, the structure of the curriculum—both what’s offered and how students are assigned to courses; and second, students’ own decisions, motivation, ability and performance (Sorensen & Hallinan, 1977). Thus, unlike research which frames course-taking purely within a “student choice” model, this study incorporates a model which recognizes both that students can only participate in courses that are available and that their course-taking decisions will be shaped by their school’s curricular structure. For example, if schools offer only a constrained curriculum composed predominantly of rigorous, academic courses, then it is likely that these are the courses that students will take (Lee et al., 1998). Thus, in the next two sections I will discuss the curriculum from both the school’s and the student’s perspectives.

**Determinants of Schools’ Curricular Structure**

A school’s curriculum and its structure are the primary determinants of what information students receive in school (Hallinan, 1987). Importantly, the exact same student might experience vastly different educational experiences in different schools with different curricular structures (Sorensen, 1987). For example, Garet and DeLany (1988) found that, controlling for students’ own achievement levels, the probability of enrolling in advanced science and mathematics courses varied systematically among the four California high schools they studied. The curricular structure—the number of levels of differentiated ability and the number of course sections at each level—influenced students’ exposure to subject matter.

What are the determinants of schools’ course offerings? At a minimum, schools are likely to provide the courses required for students to meet state and district graduation requirements. But beyond the content mandated by graduation requirements, school personnel retain the power to
decide how this content will be organized for transmission to their students and what additional content will be provided (Hallinan, 1987). This involves decisions regarding both which courses will be offered and how many and which levels will be offered. Will the curriculum be differentiated into clear academic, general, and vocational tracks? Will honors and AP courses be available? Will remedial courses be available? Will a Calculus course be offered or will the math sequence end at Pre-Calculus? How will students be assigned to courses? Different schools answer these questions differently and consequently students are exposed to different knowledge depending both on what school they attend and which track they have been assigned to (if tracks exist) within their school. School size is clearly an important determinant of how many courses schools can offer but it is just as plainly not the only factor influencing curricular structure and content. The literature on this topic is limited, but hypothesized determinants of curricular offerings also include schools’ pedagogic mission, instructional capacity such as the number and qualifications of teachers, institutional pressures, and the composition of the student body. Of these factors, I focus on school size, student body composition, and institutional pressures.

School size. The economic efficiency stream of research on school size (discussed earlier) confirmed that larger schools can fiscally justify offering more courses overall and are also more likely to offer advanced courses (Haller et al., 1990; Monk & Haller, 1993). A more recent study of the determinants of Advanced Placement and International Baccalaureate course offerings in Florida schools found that this pattern persists today. Strikingly, the authors found that only 3% of schools in the smallest size decile (enrollments of 22 to 365) offer AP/IB math compared to 100% of schools in the top three deciles (enrollments of 2,224 and above; Iatrola, Conger, & Long, 2011). However extant research also suggests that if small schools follow the Catholic school model, they will focus their resources on core academic courses and eliminate lower-level and remedial courses (Lee & Bryk, 1988; Lee, Croninger, & Smith, 1997).
These two theories are not mutually exclusive. Small schools may be likely to target their courses to middle ability students, forgoing course offerings on both the high and low ends of the spectrum. In this vein, one possible explanation for why size is related to the type as well as the number of course offerings is that large schools are simply more likely to have a critical mass of students with high, middle, and low prior achievement levels and thus can justify differentiating the curriculum to meet the perceived needs of their students (Monk & Haller, 1993; Iatrola et al., 2011). Small schools may tend to enroll more homogenous student bodies, on average, than large schools. Some research indicates that small schools have a higher variance in academic performance than large schools and thus are overrepresented in both the highest performing schools and the lowest performing schools (Wainer & Zwerling, 2006).

Another potential explanation for why large schools offer more advanced courses is that they are more likely to have teachers with sufficient qualifications and preparation to teach them, or that larger schools have more capacity to allow teachers to specialize in ways that facilitate offering a wide range of courses (Lee et al., 2000; Loveless, 1999). However, this instructional capacity argument has not been borne out by recent research (Iatrola et al., 2011) and does not convincingly predict the curricular focus of small schools, which may be less differentiated by necessity but could hypothetically have a range of curricular focuses or levels of academic intensity.

**Student body composition.** An additional determinant of schools’ course offerings are the needs and demands (real or perceived) of the student body. This is not a new concept, recall from the historical narrative of the US curriculum that many education historians believe that the original differentiation of the curriculum occurred in response to the perceived needs of an increasingly socio-economic and ethnically diverse secondary school population. For instance, there is evidence that schools with growing shares of minority students are more likely to create separate curricular tracks for high and low achievers (Lucas & Berends, 2002). Descriptively, we have known for a while that
not all schools provide equal opportunities to access advanced curriculum and that schools serving low-income and minority students are the least likely to provide these opportunities. For example, Adelman (2006) found that schools serving high proportions of low-SES or Latino students were less likely than other schools to offer math courses beyond Algebra II in difficulty. However, this is in contrast to the findings by Riehl, Pallas, & Natriello (1999), who found that schools serving low-income and minority students go to great lengths to offer advanced courses, even when there are few students qualified to take them (discussed more below).

Furthermore, a study comparing the proportion of student class time allocated to non-core, core, and advanced-core courses in New York state before and after the state increased curricular requirements found that while the share of student class time devoted to core and advanced-core course rose across the state in the years after implementation of the plan, schools in New York City with larger minority populations continued to have smaller shares of class time devoted to advanced study (Alexander, 2002). Small schools on the other hand, were found to have a smaller share of courses allocated to the core than large schools in the majority of the state but a larger portion devoted to the core in the four largest urban school districts. This points to the interaction between school size, institutional pressures, and the context of schooling.

Moreover, in the Iatrola et al. study (2011), the characteristics of the student body—particularly prior achievement—were the strongest determinants of whether a school offered AP or IB courses. The authors conclude that having a critical mass of “far above average” achieving students is the primary reason why larger schools seem to offer more advanced courses. In terms of student demand for advanced courses, achievement history seems to trump race. After controlling for prior achievement, several studies have found greater demand for advanced courses among underrepresented minority groups (Conger, Long & Iatrola, 2009; Attewell & Domina, 2008; Kelly, 2009). When applied to the context of small school reform in New York City, where small schools
were targeted to replace the worst-performing large high schools serving predominantly
disadvantaged and low-achieving students, extant research suggests that these schools will be less
likely to offer advanced courses both because of their small size and because of their academically
underprepared student body.

**Institutional pressures.** Finally, institutional pressures may influence schools’ curricular
offerings. The most obvious example of this is state or district graduation requirements. Schools
must necessarily offer the courses needed to fulfill core credit requirements and to prepare their
students for any exit exams that they must pass in order to graduate. But beyond mandates such as
these, the institutional environment can influence schools through more subtle pressures. As
discussed in greater detail below, the New York City Department of Education is in the process of
incorporating a “College Preparatory Course Index” and a “College Readiness Index,” to their
school report cards. Although the district is not yet holding schools accountable for these measures,
they are clearly promoting advanced curricular offerings.

**How Students Experience the Curriculum**

While the primary focus of this study is schools’ curricular structures and course offerings, it
is important to acknowledge that this is only half of the picture. How far students get through
course sequences is also determined by what they bring to the equation and how their own
characteristics interact with the curricular structure in place in their school. As adolescents navigate
their schools’ curricular structure, which links their prior learning and skill development to their
future educational attainment, their personal traits, social capital, and past experiences come together
to shape how their short-term decisions match up with their long-term goals (Pallas, 2003).

Prior academic experiences are particularly important in math because of the hierarchical
nature of math course-taking and the accompanying structure of prerequisites create a positional
advantage for those students who enter high school having already taken the early math courses (Stevenson et al., 1994; Schneider et al., 1998; Riegle-Crumb, 2006). In contrast to other subjects such as English or History, math offers a less “open” path, so that where students begin the sequence is strongly related to how far they progress by the end of high school (Lucas 1999). For example, a student who begins high school taking Algebra or Geometry rather than Pre-Algebra or remedial math has significant advantages for advancing farther through the math curriculum in high school. Students who do not begin high school taking Algebra I or higher have little chance of reaching advanced courses such as Trigonometry and Calculus. This is especially troubling because it is precisely these advanced courses (beyond Algebra II in difficulty) that most strongly predict college attendance (Adelman 1999, 2006; Schneider 2003). Clearly, which course a student takes in 9th grade and how successful they are in that course is critical in determining their eventual educational trajectory. Thus, how well small schools in New York City do in getting their students through the math pipeline will partially depend on which courses they have taken and passed in middle school. This fact highlights the importance of researchers evaluating small schools to ensure that they are judging small-school students against an appropriate comparison group of students with similar educational backgrounds.

Yet while starting at a higher position in the math sequence is necessary for reaching the most advanced courses by the end of high school, it is not a guarantee. Depending on schools’ curricular structures, students may confront numerous points at which decision making—active or passive, by the student or someone else—about the next move is required, with each move cumulatively shaping students’ long-term educational trajectories (Morgan, 2005; Crosnoe & Schneider, 2010). It is at these junction points that students’ racial/ethnic and socio-economic background can influence their opportunity to learn in high school (Crosnoe & Schneider, 2010). Certain students may be at a higher risk of losing an initial course advantage than others. For
example, researchers have found evidence that black, Latino, and low-SES students may be more likely to “fall out” of an initial position of advantage than their white and Asian, higher-SES peers (Hallinan, 1996; Lucas, 1999; Lucas & Berends, 2002; Oakes, 1990; Crosnoe & Huston, 2007; Gamoran & Hannigan, 2000; Riegel-Crumb, 2006).

If small schools do implement a constrained, rigorous, academic curriculum, they would facilitate disadvantaged students’ progression through curricular pipelines by eliminating some of the junction points that allow for unequal socio-economic backgrounds to manifest themselves. Students in small schools have fewer choices, and if the only courses available are both academic and rigorous in nature then they will be more likely to take a college preparatory course sequence than their counterparts in large, highly differentiated schools. However, the reverse is also true. If small schools serve a student body that is, on average, academically underprepared, most of whom have not taken and passed Algebra in middle school, the school administrators might rationally decide not to offer a course like Calculus because not enough students will have the necessary pre-requisites to take it and those resources could be better spent offering a support course for students struggling with Algebra. However, a high achieving student entering such a school, having already passed Algebra, might be “constrained” in a negative way by the limited advanced course offerings. In this manner, an individual student’s attributes interact with school characteristics such as school size, the overall student body composition, and the structure of the curriculum to shape her educational trajectory. In other words, the same school, with the same curricular structure could simultaneously improve some students’ educational trajectories while harming the trajectories of others.

Finally, while not the focus of this study, students’ social adaptation to high school can also impact their progression through course sequences. Research suggests that students feel less anonymous and more engaged in smaller schools and have better attendance rates (Chambers, 1981; Fowler & Walberg, 1991; Newmann, 1981; Lindsay, 1982; Pittman & Haughwout, 1987; Gladden,
Hypothetically this could translate to students expending more effort academically and consequently failing fewer courses. Particularly, in hierarchically sequenced subjects such as math, where failing one course affects one's ability to take and pass the next course in the sequence, fewer failures could easily translate to students progressing farther through subject pipelines. It will also be rarer for students to “fall through the cracks”—since teachers in small schools are more likely to know all students, they may be less likely to allow a bright student to take a class that is below her ability.

A Conceptual Model for Studying School Size and Student Course-taking

In developing my conceptual model for understanding the link between small school reform and students’ progression through the math curriculum, I have drawn on insights from both the historical evolution of the relationship between school size, the curriculum, and disadvantaged students’ access to high-status knowledge and the existing literature on small schools in which the reform was embedded (and the shortcomings of applying this literature to the current context). I have used these insights to augment the constrained curriculum hypothesis, resulting in a more complex model for understanding the link between school size and student course-taking. As conceptualized in my model, policy effects of small school reform can flow either directly, through schools’ curricular structures: the courses, content, and rigor that students experience, or indirectly through student engagement (see Figure 3.1). This study will focus primarily on the direct pathway from school size, through curricular structures, to student course-taking outcomes (denoted by solid lines) and not on the pathways that flow through student engagement (denoted by dotted lines).

How far a student progresses in math by the end of high school will depend on the structure of her school’s curriculum (which courses it offers and requires, how students are assigned to courses) and her own social and academic background characteristics that could influence her placement in the curriculum (particularly prior achievement and 8th grade math course). The courses
that students take in ninth grade, combined with their response to instruction in those courses shape not only their learning in ninth grade, but also their curricular trajectory throughout high school. If small schools constrain their curriculum *upwards* to an academically intense level of difficulty, there will be little variation in course-taking because all students will be compelled to take an academic, college preparatory sequence of courses. This should result in small-school students progressing farther through subject pipelines and achieving higher on Regents exams than their counterparts in large schools, particularly students who would otherwise take less-demanding courses. It also should result in a more equitable social distribution of student course-taking by race and class in small schools. Following this logic, students in small schools should ultimately exit high school more prepared for college and the workforce than students in large schools.

Yet, as shown in Figure 3.1, the success of small school reform is conceptualized here to depend on a number of mediating factors. Student body composition is also hypothesized to impact schools’ curricular structures. Schools with large concentrations of academically underprepared students may implement less rigorous curricular structures with fewer advanced course offerings. This relationship may be especially pronounced in small schools, which lack the capacity to offer courses both at the remedial and advanced ends of the spectrum. Schools with many underprepared students may need to answer standards-based institutional mandates by offering extra remedial or support courses in place of advanced courses to help their students pass the required exit exams.

The number and quality of the teaching staff may also influence how many and which courses are offered. Small schools will necessarily have fewer teachers but their teaching staffs may also differ with respect to qualifications. If small schools serve more disadvantaged students, for example, they may attract less qualified teachers. Conversely, the novelty of a new reform coupled with policies providing the new small schools with more autonomy over staffing may attract higher quality, more energetic teachers.
Individual student characteristics are also expected to interact with school size and the related curricular structure to influence students’ educational trajectories. For example, a student with low entering academic achievement is hypothesized to profit more from a constrained curricular structure than a student with high entering achievement. Demographic characteristics such as race and class are hypothesized to matter more for students’ course-taking in large schools with comprehensive curriculums. Finally, the institutional context is also conceptualized as influencing schools’ curricular structure, and thus students’ educational experiences. Since all public schools in New York City—small and large—experience the same institutional environment (discussed in detail below), this may serve as a homogenizing influence when it comes to schools’ decisions regarding their curricular structure. Institutional pressure, may also lead to disproportionate course-label inflation in schools with high concentrations of academically underprepared students.
*For brevity I use “Intensity” here to refer both to which course in the sequence (i.e. Algebra vs. Calculus) and its level (i.e. regular vs. AP).

Figure 3.1. Conceptual model of how school size influences student course-taking and educational outcomes in math.

The Threat of Course-Label Inflation in New York City Schools

A consistent dilemma facing research on student course-taking is the concern that course titles may not accurately portray course content or course rigor. Even more concerning, is the possibility that some schools might be less accurate in their designations than others and that this type of course label “inaccuracy,” or more likely “inflation,” might vary systematically across high schools based on characteristics such as student body composition, and in this case, school size. Thus in this section I will describe in more detail the concept of course-label inflation and discuss
the elements of the New York City institutional context, which might pressure some schools to
over-state the level or rigor of their courses.

**Course-label Inflation.** Although schools are increasingly being evaluated on their students’
performance on exams, there are still institutional pressures for schools to signal their legitimacy as
educational institutions through other channels such as course offerings and students’ credit
accumulation. State or school district mandated course credit requirements are a good example of
this. Furthermore, as state and federal pressures to “upgrade” the curriculum become more and
more intense, and as the belief in the need for all students to follow a college preparatory curriculum
becomes institutionalized, schools are more likely than ever before to feel pressure to project the
image of a strong academic curriculum, regardless of whether they have one or not (Dougerty,
Mellor, and Jian, 2006a; Mathews, 2006; Cogan et al., 2001). Dougherty, Mellor, and Jian (2006a)
caution that “course credit inflation” is becoming an increasing problem. Similar to the concept of
grade inflation, in which the knowledge and skill of the average student receiving an “A” declines
over time, Dougherty and colleagues believe that the level of content mastery for the average
student receiving credit for a course with a given title may be declining over time.

Symptoms of this unintended consequence have been evident since the advent of the
standards and accountability movement in the 1980s. For example, research indicates that initial
state graduation requirements for a certain number of years/credits per subject were creatively
circumvented by schools—math course-taking did indeed increase, but all of the gains were in math
courses below algebra in difficulty. Some high school transcripts applied the label "pre-calculus" to
any math course before calculus, according to one study, and some students who had taken "pre-
calculus" as indicated by their transcripts, had skills so rudimentary that they were forced to take
basic algebra in their first year of college (Mathews, 2006).
Over the past decade, researchers and policy-makers have become increasingly worried about the disconnect between reported increases in the percent of students taking advanced courses and national test scores, which have essentially flat-lined for the past four decades. A recent federal study of nearly 38,000 high school transcripts showed that the proportion of graduates completing a rigorous curriculum rose to 13 percent in 2009 from 5 percent in 1990 (Nord, Roey, Perkins, Lyons, Lemanski, Brown, & Schuknecht, 2011). By contrast, the 2009 results of the federal test that measures change in achievement levels over decades showed that the nation’s 17-year-olds were scoring no higher in reading and math than in 1973. SAT scores have also dropped or held steady since 2000 (Dillon, 2011).

Researchers at Michigan State University examined the test scores of 13,000 American eighth-grade students who participated in an international math and science exam known as TIMSS. They compared the schools’ math courses — ranging from remedial through “enriched” — with the content of the textbooks used in them. There was a mismatch between the course title and the textbook employed for nearly 30 percent of eighth graders. In about 15 percent of the cases, the textbook covered less advanced areas of math than the course name suggested (Cogan et al., 2001).

And it is not just academics who are alarmed about the possibility of course credit inflation. In 2005, the College Board was sufficiently concerned about the quality and rigor of courses claiming to be “AP” that they initiated an audit process to bolster the credibility of the AP brand name and assure college admissions officers that the rapid expansion of students taking AP courses had not led to a general watering down of the AP curriculum (Honawar, 2005). The results of the audit were interesting: of the 134,000 syllabuses for AP courses that were submitted by teachers, only 67 percent were approved on the first review (Jaschik, 2007). Furthermore, 2,081 high schools that had been offering AP courses declined to submit to the audit after they reviewed the standards that would be applied. Even though these high schools offered an average of 5 AP courses each,
they realized that the curriculum in their courses were sufficiently far below the official AP standards that it wasn't worth going through with the audit (Jaschik, 2007). And perhaps most tellingly, there were clear differences in the quality of AP offerings at high schools serving affluent students as compared to high schools serving students of lower socio-economic status. Only 3 percent of the wealthiest high schools had to improve textbooks as part of the audit process, while 22 percent of low-income high schools were found to be using deficient texts. Similarly, low-income high schools had to increase laboratory time at twice the rate of their wealthier counterparts (Jaschik, 2007).

**Diagnosing Course-label Inflation.** One way to investigate whether course credit inflation does indeed exist is to test whether taking advanced courses is a good predictor of later educational outcomes such as college matriculation and completion. For example Geiser and Santilices (2004) found that students’ number of advanced course credits did not predict their likelihood of persistence in college, but that scores on the AP and SAT II curriculum-based exams did. In the same vein, Dougherty et al (2006b) found that a high school’s percent of students taking and passing AP exams was a much better predictor of college graduation rates than the percent of students taking AP courses but not passing exams. Another, perhaps more direct strategy for unearthing course credit inflation is to examine the relationships between course completion and subsequent performance on more objective, standardized tests that measure the same content. For example, Dougherty (2006a) reports that, in Texas, 60% of low-income students, 65% percent of African-American students, and 57% of Hispanic students who received credit for Geometry and Algebra 2, failed a state exit exam covering Geometry and Algebra 1. The corresponding percentages for non-low-income and white students were 36% and 32% respectively.

Dougherty et al (2006b) also question the extent to which academic research based on curricular data from the 1980s and early 90s—essentially all data using High School and Beyond (HS&B) or the National Educational Longitudinal Study (NELS)—remains valid today. The HS&B
and NELS student cohorts attended high school in an era when there was as yet little pressure for all students to take advanced course-work and consequently only the most prepared students did so. Dougherty argues that in this era, there was little pressure to water down course content, because most students following the college preparatory curriculum were academically prepared to master the appropriate content. The implication is that current research on course-taking patterns and curricular rigor cannot afford to rely on the same assumptions as studies from this earlier era. We can no longer blithely assume that course titles are accurate representations of course content or that students who have received credit for a course have necessarily demonstrated mastery over its subject matter. Therefore, in this study, I heed the caution of Dougherty et al. and the institutional theorists and examine the extent to which course-label inflation might be at play in New York City schools.

**Interaction Between Institutional Pressure, Stratification, and Course-label Inflation.**

How will small high schools in New York City respond to the institutional pressures generated by state mandates and high-stakes exit exams? Some research suggests that both their disadvantaged student population and their small size may increase the likelihood of practicing course-label inflation. The better researched of these two interactions is the relationship between institutional pressure, schools’ student body composition, and the validity of their course labels. Within our highly stratified education system, state mandates are clearly less easily met in some schools than others. Research indicates that high schools with concentrations of underprepared students may feel the pressure to project the image of a more rigorous curriculum than they actually have.

As Mary Metz (1989) discovered in her research in both economically advantaged and disadvantaged schools, all schools tend to follow a “common script”, particularly when it comes to curricular decisions, and the more uncertain the environment the more tightly they will cling to the script. Metz found that in the poor, minority school she studied teachers did not attempt to change
their pedagogical approaches despite the fact that current organizational practices and teaching strategies clearly weren’t working. Instead, their response was to “water down” the common curriculum and alter their expectations. This was not technically efficient but maintained their sense of “legitimacy”. Research also indicates that schools serving poor and minority students go to great lengths to offer advanced courses, even when few of their students have sufficient academic preparation to succeed in them, because they prize legitimacy over efficiency (Riehl, Pallas, & Natriello, 1999). Metz believed that the standardization of the common script across schools serves a symbolic purpose; it covers up educational inequities and allows society to believe in the existence of equality of educational opportunity and meritocracy. In evaluating the effects of urban small school reform on curricular offerings, it is important to keep the concept of a “common script” in mind and to evaluate whether “Algebra 1” in one school is equivalent in academic rigor to “Algebra 1” in another school.

Institutional pressures in New York. New York State has been at the leading edge of ramping up the institutional pressures on schools and has a long history of administering high school examinations (Debray, 2005). There has been a course of study leading to a Regents diploma since 1878 (Bishop, 1998). However, prior to 1996 students earned either a Regents diploma, which required a college preparatory course of study and passing the associated examinations, or a local diploma received by the remainder of students after passing minimum competency tests (Debray, 2005). In 1996 the Board of Regents passed the Regents Action Plan, mandating the phasing out of the competency tests, and instituted a new policy whereby all students would be required to pass Regents examinations in five subjects: United States history, global studies, mathematics, English/language arts, and science. They also increased the number of core courses required to graduate from 20.5 to 22. Starting in 1996, students in New York City must pass all five exams with a 65 or better to earn a Regents diploma and a 55 or better to earn a local diploma. The testing
requirements are now in their final stages of being phased in; students who entered high school in 2008 and were on track to graduate in 2012 no longer had the option of a local diploma and must pass the five exams with a 65 or better to graduate at all. An Advanced Regents diploma requires passing 7-9 Regent exams with a 65 or better.

For context, roughly 65% of all students in the 2006 cohort graduated in four years, and nearly 14% of those graduates attained only a local diploma. Furthermore, using data collected by state and community colleges, testing experts on a state committee determined that a 75 on the English Regents and an 80 on the math Regents roughly predicted that students would get at least a C in a college-level course in the same subject. Scores below that meant students had to often take remediation classes before they could do college-level work (Otterman, 2011). Disturbingly, only a fifth of the 2006 cohort of students, who were on track to graduate in 2010, met this elevated criterion. Amid mounting concern regarding the college readiness of its graduates, New York City is in the process of incorporating a “College Preparatory Course Index,” a “College Readiness Index,” and a “College Enrollment Rate” into the Progress Report Card for high schools. These metrics were incorporated, but not scored, in the 2010-11 Progress Report, and were scored in 2011-12 (NYC DOE, 2011a).

Clearly, schools in New York City are under enormous pressure to upgrade their curriculums and ensure that all students take the courses needed to prepare them for the high-stakes Regents exams. Interestingly, the fact that all schools in New York City, regardless of size, are feeling these same pressures could theoretically reduce the positive course-taking effects we might expect to see in small schools due to the “constrained curriculum” hypothesis. Even large schools may feel pressure to reduce the number of low-level courses that do not help students pass their required Regents Exams. However, the pressure put on New York City schools to get their students through core academic courses, could also lead to systematic course-label inflation among schools serving
students with weak academic skills. Since small school reform in New York City targets students in disadvantaged areas, it is possible that they may disproportionately succumb to the pressure to practice course-label inflation. To address this concern, my study will assess the extent to which success (as measured by course grade) in a given course is associated with success on the relevant NY State Regents Exam for that course.

**Summary**

In this chapter I have argued that a new, more complex theoretical model is needed to frame the relationship between small school reform and student course-taking that not only conceptualizes the direct pathways through which school size might impact student course-taking but also includes student and school factors that might mediate this relationship—particularly schools’ student body composition. The model I have laid out in this chapter builds on the constrained curriculum hypothesis, which was formulated to explain the positive findings from research on school size at a national level, to make it more applicable to the contemporary, urban, small high school reform. This model has helped me to develop analyses that can capture the complexity of the relationship between small school reform, schools’ curricular structures, and student course-taking. In the next chapter, I will discuss the data and methods that I utilize to address my research questions. I will also describe the location for my study, the New York City high school system, in greater detail.
Chapter 4
Methodology

In structuring my study on the relationship between urban small school reform in New York City and student course-taking, I have drawn on three important guides: first the historical evolution of the relationship between school size, the curriculum, and disadvantaged students’ access to high-status knowledge; second, the existing literature on school size and small school reform (and the shortcomings of this literature); and finally a conceptual model which builds upon the constrained curriculum hypothesis to describe the mechanisms through which small school reform might impact students’ progression through the math curriculum. I have drawn on these three guides to develop an analytic strategy that best addresses my research questions.

My study focuses on three central questions surrounding the curricular outcomes of small school reform in New York City. First, I descriptively examine the curricular opportunities that small schools provide for their students. To what extent do small high schools implement different curricular structures and provide their students with different course offerings in mathematics than their larger counterparts? Second, I investigate the impact that small schools have had on student course-taking. Do students in small schools progress farther through math curricular pipelines than similar students in mid-sized or large schools? And finally I address the rigor of courses across schools. Are courses in the new small schools comparable to those in larger schools in terms of their academic standards? In this chapter I discuss these questions in greater detail and outline the data and methods that I will employ to answer them. However before diving into the methodological details it is important to briefly discuss the context of small school reform in New York City.
Context of Small School Reform in New York City and Implications for Analyses

The New York City public school system is the largest and most complex in the United States, serving roughly 1.1 million children in nearly 1,700 schools (NYC DOE, 2011b). The complexity of the system creates both challenges and opportunities for evaluation. In this section, I will briefly discuss the relevant aspects of New York City education policy, particularly the changing landscape of high school options during the first decade of the 21st century. In assessing the impact of small school reform, it is crucial to establish both what counts as a “small school” for the purposes of the analyses and also to identify the appropriate “counterfactual”—that is, what would have happened in the absence of the reform. To do so, it is necessary to first understand the full range of high school options as well as the school system’s assignment mechanism.

Layers of Small School Reform

As discussed in Chapter 2, over the past 50 years New York City has experimented with several waves of small school reform, each with different leaders, goals, and assumptions about what the reduction in school size would accomplish. In the 2002-03 school year there were 58 small high schools in NYC remaining from previous rounds of small school creation; six years later there were a total of 161 small high schools (Quint et al., 2010). Thus, while the focus of this study is on the most recent wave of small school reform, beginning in the 2002-03 school year, it is important to recognize that some “veteran” small high schools remain from earlier waves of reform (Quint et al., 2010). Moreover, the new small schools are distinct in that they received start-up support as part of the reform as well as increased autonomy over decisions like hiring, and temporary exemptions from

---

7 The DOE counts schools as “small” if they enroll 550 students or less, with a maximum 9th grade enrollment of 175 (MDRC, 2010). This is noticeably smaller than the 600-900 range advocated by Lee & Smith (1997).
serving special education students and English language learners during their first two years of operation (Quint et al., 2010).

Children First Initiative

Creating an educational marketplace. Recall from the literature review, that small school creation under the Bloomberg/Klein regime was introduced as part of a more extensive policy initiative known as Children First. The replacement of large, dysfunctional high schools with small schools was part of a broader push to expand the range of high school choices available to all students and to facilitate the creation of an educational “marketplace” in New York City (Hemphill & Nauer, 2009). Prior to Children First, eighth graders in low-income areas were generally assigned to their neighborhood, zoned high school. School choice did exist—students could apply to up to five schools that they were interested in attending other than their zoned school—however, this option was highly skewed in favor of more advantaged, higher ability students. First, higher SES students were more likely to be aware that choice was an option and second, many schools had selective admissions processes and would deny access to students who did not meet certain academic criteria. School principals essentially had the power, in most instances, to deny admittance to students who did not meet their standards. The new system under Children First forces all eight graders to rank their top twelve high school choices. Students learn about schools through their guidance counselors, social networks, high school fairs, and the 300 plus page New York City High School Directory. The fact that students have an element of choice in where they attend high school underscores the need for any research on high school effects in New York City to address student selection into schools. In the context of this study, the concern is that students who choose small schools may have different observed and unobserved characteristics from students who choose larger schools and that these characteristics might also be related to their persistence through the
math curriculum. Student selection into schools can generate selection bias at the school level as well, as schools may tailor their curricula to the real or perceived needs of their student bodies.

**High school options: a range of size and selectivity.** Eighth graders’ decision of which high schools to choose and in what order is complicated by the fact that there are many different types of high schools, which vary in their selectivity, governance, grade configurations, and targeted student populations.

**Types of high schools:8**

- *Charter schools* are publically funded schools that are governed individually by not-for-profit boards of trustees. They are operated independently of the DOE but must still meet educational standards set by New York State.

- *District 75 schools* serve only full-time special education students. District 75 consists of school organizations, home and hospital instruction and vision and hearing services.

- *General high schools* are traditional high schools, operated by the DOE, serving students in grades 9-12.

- *Middle/high schools* typically serve grades 6-12 or 7-12. They are intended to ease students’ transition from middle to high school by allowing them to maintain relationships with teachers and staff and remain in a familiar setting.

- *Transfer schools* are full-time schools targeting students who are overage and undercredited. They are generally small—some of them former alternative small schools created as early as the 1970s—and are designed to help students overcome obstacles to graduation.

Even at its broadest, this study will only examine NYC DOE operated high schools serving first-time 9th graders. It will not include charter schools, District 75 schools, or Transfer schools (sample selection discussed in greater detail in the Data section). The new, small high schools, that constitute the treatment group for this study, are all either general or middle/high schools. Beyond school type, many schools, particularly the mid-size and large high schools, contain multiple programs, all of which may differ in their admissions criteria.

---

8 Information on high school types and eligibility criteria taken from Quint et al. (2010) and the New York City Department of Education website: http://schools.nyc.gov/default.htm.
School/Program Eligibility Criteria:

**Academically selective**
- *Specialized high schools* are the most selective high school option, only serving students who have demonstrated academic and/or artistic excellence. Eight require a competitive score on the Specialized High School Admissions Test (SHSAT), and the ninth admits students based on an audition and a review of academic records.

- *Screened* programs select students based on whether they meet specific criteria. These criteria are generally academic; schools select students on the basis of having scores that meet a certain threshold on district tests administered in seventh grade and then rank students in order of preference. These thresholds vary from program to program. Screened programs can also require students to attend an interview or submit a portfolio of work.

- *Audition* programs admit students based on an audition or portfolio.

- *Educational option* programs are designed to create a mixed-ability student body. They admit students based on a combination of ranking and random selection by computer (16% each from the top and bottom achievement categories and 68% from the middle).

**Academically non-selective**
- *Limited unscreened* is a selection method introduced with the new small schools. Limited unscreened programs do not give preference based on any academic criteria. Instead, they give preference to students who (1) live within a certain geographic area, and (2) have attended the school’s open house, or their booth at a school fair, or are otherwise “known” to the school. Almost all of the new small schools are single-program, limited unscreened schools. They were intentionally designed to be academically non-selective because they were targeted to replace the large, dysfunctional, zoned schools that were closed.

- *Unscreened* programs set no admissions criteria.

- *Zoned* programs have no admissions criteria except for requiring that students live in the attendance area of the school within which the program is located (or they give preference to such students).

Parental support and family social and cultural capital can be of crucial importance for navigating this incredibly complex system. Interview data suggests that low-income students and students from non-English speaking families in particular have difficulty navigating this system, sometimes selecting a list of high school choices that is not well aligned to their needs (Hemphill & Nauer, 2009). Moreover, recent case study research indicates that some schools categorized as
“unscreens” actively circumvented district rules against selecting students based on performance in order to recruit and retain a population that would meet local accountability targets (Jennings, 2010).

In 2007, eighth graders could apply to 595 different high school programs (Pallas & Riehl, 2007). Once students apply, the DOE uses a centralized High School Application Processing System (HSAPS), which is a computer-based algorithm that matches students to schools while simultaneously taking into account students’ preferences for schools and schools’ preferences for students. The process was modeled after the matching system that places medical students with residency hospitals. It was introduced in the 2003-04 school year and has been adjusted every year to make it work better. By the 2007-08 school year, HSAPS successfully placed 50 percent of students in their first-choice school and 80 percent in one of their top three choices. There are three rounds of matching, students who are not offered a seat in the first round can reapply to high schools with open seats in the second round, and so forth. Finally, each year a substantial portion of students—about 12.7 percent of incoming freshmen—are assigned to a high school after the conclusion of the HSAPS process (Quint et al., 2010). These “over-the-counter” students (in DOE parlance) may be recent transfers into the district, may have recently switched from private to public school, or may have moved to a neighborhood unreasonably far from the school to which they were initially assigned. These students are usually placed in any school near their home that has an opening (Quint et al., 2010).

In terms of demand, schools in Manhattan, programs with higher average achievement, and schools with fewer black and ELL students are in higher demand than other programs (Pallas & Riehl, 2007). However, the single biggest predictor of how popular a school would be in any given year was the popularity of that same school the year before (Pallas & Riehl, 2007). All in all, the high school choice system in New York City, which involves schools choosing students as well as students choosing schools, necessitates the adoption of an analytic strategy that can address
selection bias. At its broadest, this study will describe both the enrollment patterns and the curricular structures of NYC DOE operated public high schools across the full range of size and selectivity. At its narrowest, it will focus on estimating the curricular effects of only the new, small, non-selective (NSN) schools for the students who attend them—the treatment effect on the treated.

Data & Methods

Data. My dissertation employs a unique, longitudinal database of New York City student and school level data from 1999-2010, constructed from New York City Department of Education administrative data, 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 from the 1999-2000 school year through the 2010-11 school year. The data include all students’ English Language Arts and Mathematics scores from grades 3-8, their full transcript files from high school (with course titles and grades), and their Regents Exam dates and scores. My study will focus on high schools, although the dataset includes information on students’ middle schools (as long as they attended a public middle school in New York City). This allows for the possibility of more complex analytic strategies such a matching technique that reduces the threat of selection bias by comparing only students who attended the same middle school but went on to attend different types of high schools.

For this study I focus on two cohorts of students, those entering ninth grade in 2005 and 2006. These are the two most recent cohorts of students whom I can follow for a full four years of high school. Since new small schools have been created every year since the beginning of the reform in 2002, selecting the most recent cohorts maximizes the sample of students attending new, small schools. It also means that the schools that opened in the early years of the reform have had a
chance to move beyond the “honeymoon” phase of the NYC DOE’s start-up accommodations as well as the initial burst of staff energy that can sometimes accompany the creation of a new school.

Sample. My descriptive sample includes all NYC DOE operated schools that serve first-time 9th grade students and are not in the full-time special education district. It excludes charter schools, District 75 schools, and Transfer schools (2005-06 N=322; 2006-07 N=340). For descriptive statistics on schools’ student body composition, all students entering 9th grade in 2005-06 or 2006-07 are included (n=132,164). However, when I begin to investigate schools’ curricular structures and student course-taking, I further exclude students who transfer schools during high school, whether they leave the New York City public school system entirely (roughly 13.7 percent of the base sample) or simply transfer among school within the system (an additional 12.5 percent). Transfer students pose a problem because their accomplishments are not attributable to one school alone. This is particularly a concern for my outcome of interest, progress through the math curriculum, because it is measured after the completion of a full four years of high school. Because my analyses do not include transfer students, my results are only generalizable to students who remain in the same school for the duration of their high school career. Students who drop out of high school are retained in the sample so long as they only attended one high school prior to dropping out (see below for a more detailed discussion of drop-outs).

For research question 2, which focuses on the impact of small schools on students’ progression through the math pipeline, the sample of students is restricted to those who attended a New York City public school for eighth grade (n=88,241). This restriction is necessary for my propensity score matching strategies, which at a minimum match students based, in part, on the characteristics of their middle schools, and at the maximum force an exact match on students’ middle schools (discussed in more detail in the Analytic Approach section). Eighth grade information is also needed to control for students’ academic achievement prior to entering high
school. For research question 3—which assesses the rigor of courses across schools by estimating the differential between performance in specific courses and achievement on the subsequent Regents exams—the sample is necessarily limited to students who have taken the relevant math Regents Exams (Math A, n=70,222; Math B, n=22,570).

**Measures**

_School size: how small is small?_ One weakness of the literature on small schools is the lack of consistency regarding what actually constitutes a “small school” (Darling-Hammond, Ross, & Milliken, 2006). For example, educator and author, John Goodlad (1984) claimed that he “would not want to face the challenge of justifying a senior high school of more than 500 to 600 students” (p. 310). By contrast, Valerie Lee and Julia Smith (1997) concluded from their research that, while schools under 600 students were the most equitable, the sweet spot for maximizing both performance and equity was a mid-sized high school enrolling between 600-900 students. High schools of this size, they argued, were large enough to provide a solid curriculum, yet small enough to limit curricular differentiation. Finally, the New York City Department of Education has chosen to set an enrollment of 550 students as the upper limit for their definition of a small school.

MDRC’s reports on small schools in New York City (Quint et al., 2010; Bloom et al., 2010) align their school size categories with the NYC DOE’s designations. They define schools enrolling 550 or fewer students to be small, 551-1,400 as mid-sized, and over 1,400 as large. Schwartz, Stiefel & Wiswall (2013) also adopt the 550 student cut-off to examine small schools in New York City.

Thus, to incorporate a policy-relevant figure and to allow for comparability with other studies evaluating small schools in New York City, I adopt the same size classifications as the extant literature (small=fewer than 550, midsize=551-1,400, large=over 1,400). Since some of the schools in my sample are middle-high schools, I utilize only the 9th through 12th grade enrollments when calculating school size. Furthermore, because some of the schools in my analyses are still in the
process of “phasing in” (adding one grade per year), if a school is still in the phase-in process, I further require that the 9th grade class has a maximum enrollment of 175 students to be classified as small. If schools differed in their size classifications across cohorts due to minor fluctuations in their student enrollment across these cut-points, I maintained their size classification from the first cohort (2005-06).

School Academic Selectivity. 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, I 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 is greater than or equal to 50 percent, the school as a whole is considered to be academically selective for that year. Information on program selectivity and number of seats is taken from the New York City High School Directory from the prior school year. In some instances, middle-high schools did not provide exact information on the number of selective seats because they give preference to students who are already attending the middle grades of their school. In these cases I calculated the average entering achievement based on students’ eighth grade math and ELA test scores of both selective and nonselective schools in the rest of my school sample (excluding the specialized high schools) and utilized these as guidelines for classifying the remaining middle-high schools.

New vs. Veteran Schools. Because this study focuses on the most recent wave of small school reform in New York City, which began under Mayor Bloomberg during the 2002-03 school year, I distinguish between “new” and “veteran” small schools. Schools are characterized as “new” if they
were founded during or after the 2002-03 school year and as “veteran” if they were established earlier (New=1; Veteran=0).

Math pipeline. Instead of simply calculating the number of math credits taken by students, like much of the early research on the curriculum, I employ a technique of measuring student course-taking pioneered by Lee et al. (1998) which utilizes the highest level of math passed during high school to assess how far students proceed through the math pipeline (ie. 0=no math through 8=advanced topics beyond Calculus; see Table 4.1)\(^9\). It does not matter which semester students take their most advanced course in. This technique emphasizes which courses students take during high school rather than how many years of courses they take and is more consistent with the literature on what matters both for learning during high school and for college and career readiness. Table 4.1 displays the curriculum equivalents of the levels of the math pipeline measure. I have designed the pipeline measure such that one unit approximates one year’s worth of math course-taking and 0.5 units approximates one semester. I do however separate advanced AP math courses into their own category, even though few students would take 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, I chose to designate reaching AP/IB math courses as progressing farther through the math curriculum, than reaching the regular versions of those same courses.

Over the past decade, New York City high schools have experimented with two different math curriculums. The first was an integrated approach (introduced in 1999) in which students take a three-semester sequence called “Math A” which covers the equivalent of Algebra I and the first half of Geometry, followed by another three-semester sequence called “Math B” covering roughly

\(^9\) Although “College Now” courses, which allow students in participating high schools to take courses at CUNY for either college or high school credit, are included in the transcript file provided by the NYC DOE, it is unclear whether all College Now courses completed by students are accounted for in these data.
the second half of Geometry and Algebra II/Trigonometry. The second curriculum, implemented during the 2008-09 school year, constituted a return to the more standard, non-integrated sequence, of Algebra I, Geometry, and Algebra II/Trig. For the two cohorts in my study (those beginning 9th grade in 2005 and 2006), the vast majority of students took Math A and B sequences, culminating in Math A and B Regents exams (93.8 percent of the math Regents exams were Math A or Math B for these cohorts).

Table 4.1. Curriculum Equivalents of the Math Pipeline Measure

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

*First math course attempted.* This variable measures the distance through the math pipeline of the first math course students take in high school. It utilizes the same metric as the pipeline measure (0=no math, through 8=advanced topics), however unlike the pipeline measure, it designates the first course attempted, not the first course passed. This is an interesting measure because it can be
thought of in two ways. First, it might be viewed as a control for the level of mathematics students achieved in middle school. In other words, some students matriculate to high school having already passed Algebra I, whereas others may not have even passed Pre-Algebra. High schools must work with the skills that students arrive with, therefore a school that brings students from Pre-Algebra through three years of math should be given the same amount of credit as a school that brings students from Geometry through three years of math. But on the other hand, placement in the first math course taken in high school can also be viewed as the first action that high schools take that will either help or hinder students’ progress through the math curriculum. For example, assigning incoming freshmen to remedial math might forever alter their math experience in high school, preventing them from ever reaching advanced, college preparatory courses. In other words, the first math course should also be viewed as a mediator of high schools’ effect on students’ progress through the math curriculum. For this reason, I utilize the first math course measure in my descriptive analyses, but I do not use it as a control in the propensity score matched, analytic models because it is inappropriate to include mediators (or any variables measured post-treatment) in causal analyses. Instead, I rely on the propensity score matching to select an appropriate control group for students who attend a NSN school.

Curricular structure. For my descriptive analysis of curricular structures I construct several measures of the mathematics curricular structure. Lacking access to school reports of which courses they offer, I utilize data on student course-taking to create proxy measures of schools’ curricular structures. In some ways, this strategy can be considered more advantageous because it guards against the possibility of schools listing course offerings in their manuals that do not actually exist in practice. Again drawing from the curriculum research of Lee et al. (1998), I include multiple measures of curricular structure in order to capture the average progression through the math curriculum as well as the availability of lower-level and upper-level courses:
1. The average of the highest level of math course completed in each school.
2. The variability of math course-taking in each school (the within-school standard deviation of the highest course completed).
3. The percent of total credited math courses below Algebra.
4. The percent of total credited math courses in the prep/support category.
5. A measure of whether Calculus is offered in each school. This will take the form of a dummy variable set equal to 1 if more than two students took Calculus in school \( j \) in cohort \( t \).
6. A measure of whether any math Advanced Placement (AP) or International Baccalaureate (IB) courses are offered in each school. This will take the form of a dummy variable set equal to 1 if more than two students took an AP/IB math course in school \( j \) in cohort \( t \).

Course Grades. New York City high school teachers provide their students with numeric as well as alpha grades, which range from 0-100 (however, it is common practice to assign failing students a 55 regardless of how low their course performance was). Consequently, this study employs a continuous, numeric measure of students’ grades, conditional on passing, to assess the relationship between performance in courses and performance on the subsequent Regents’ Exams. Limiting the analyses to passing students should not pose a problem for my analyses, because the sample for research question 3 is already limited to those students taking a Regents Exam in a given subject.

Regents Exams. Students in New York City’s public high schools must take statewide Regents examinations in a number of subject areas in order to receive a diploma. To receive a regular high school diploma, students in New York State must pass, with a score of 65 or higher\(^{10} \), Integrated Algebra (or Math A). To receive an Advanced Regents Diploma, students must also pass one or two additional math exams (Geometry and Algebra 2/ Trigonometry if they've taken Integrated Algebra, or a choice of Math B or Alg2/Trig if they've taken Math A). Because the vast majority of students in my sample took the Math A and (optionally) the Math B Regents, I utilize only these two exams.

\(^{10}\) For the 2005-06 and 2006-07 cohorts, a local diploma requiring only a score of 55 in five subject areas, was also available. It has since been phased out.
in my analyses. For students who take math Regents exams, I employ a continuous measure of how they perform.

Performance Differential. I use the concept of rigor to refer to the standards of performance within courses (Adelman, 1999). In other words, Algebra I can be taught in an academically challenging, or rigorous, fashion where students are held to high standards of content mastery or it can be watered down in such a way that students can pass the course without learning the appropriate skills. To assess the rigor of courses across schools I compute a measure of the discrepancy between students’ average course marks and their marks on the relevant Regents Exam. I call this gap the “performance differential.” I do this separately for Math A and Math B. To take Math A as an example, I first average the numeric grade of all courses covered on the Math A exam that students (a) passed, and (b) took in high school (I do not want to hold high schools accountable for courses taken in middle school), and (c) took prior to attempting the Math A Regents exam for the first time. I then subtract this average course mark from the students’ Math A Regents mark (only their first attempt). The result is an estimate of the gap between students’ performance in their courses and their performance on the more objective Regents exam. Zero on this measure would indicate perfect alignment between course grades and Regents performance, negative scores indicates students performed worse on the Regents than in their courses, and positive scores indicate that students performed better on the Regents than in their courses.

Both teachers’ grades and the New York State Regents exams aim to accurately signal students’ content mastery in the relevant subject, yet according to classical measurement theory, they necessarily do so imperfectly. Moreover, issues surrounding measurement error are heightened even further in the case of difference scores—such as my “performance differential”—whose reliabilities are generally weaker than either of their constituent measures. It is possible to calculate the reliability of a difference score if the reliabilities of the component measures are known. However, in the case
of my performance differential measure, the reliability of teacher grades is unknown and the
reliabilities of the NY State Regents exams are not publically available. The results from my
performance differential analyses are consequently interpreted suggestively. The estimates of
variables measured with error will be biased toward zero (Fuller, 1987), thus the relationships that I
present between school characteristics and students’ performance differentials are likely conservative
estimates of the true relationships.

*Social and academic background.* Student-level academic measures include students’ seventh and
eighth grade Mathematics and English Language Arts (ELA) scores (standardized within cohorts;
M=0, SD=1). Additional socio-demographic measures include dummy coded indicators of students’
gender (female=1, male=0), English language learner status (yes=1, no=0), free/reduced price lunch
eligibility (yes=1, no=0), Special Education status (yes=1, no=0), and whether the student was
classified as overage for eighth grade, defined as 15 months older than the New York City minimum
age for eighth grade (yes=1, no=0). I measure race/ethnicity using a series of dummy-coded
measures of whether students were Asian, Black, Hispanic, or Native American/Multi-ethnic, with
whites as the un-coded comparison category. I also account for the number of absences and tardies
students registered during eighth grade as well as student mobility, coded as the number of middle
schools a student attended.

*School characteristics.* In addition to size, selectivity, and era of founding, school characteristics
include borough, Title 1 eligibility, and aggregate student characteristics such as school racial/ethnic
composition, percent of students who are eligible for free/reduced lunch, overage, English language
learners, and special education students. They also include school-average entering achievement, and
school-average absence and tardiness rates. Moreover for all schools in my sample I have measures
of several aggregate teacher characteristics such as the percent of teachers with less than three years
of experience, with masters degrees, teaching out of certification or without certification, and the
five-year teacher turnover rates. Since this dataset is not limited to high schools, I utilize the same aggregate measures for students’ middle schools for my multilevel propensity score matching strategy (discussed below). For school size in middle school, however, I employ a continuous measure of student enrollment for grades six through eight.

**Missing Data**

Given the difficulty of collecting and entering data for the sheer number of students in the New York City public schools, many of whom are highly mobile, it is no surprise that a significant portion of students in this dataset are missing data on at least some measures. Furthermore, it is likely that the data are not missing completely at random. In general, low-income students are more likely to be missing data due to their higher rates of mobility and absences. Students whose families do not speak English may also be more likely to be missing data because their parents may be less likely to complete the administrative questionnaires used to obtain demographic data such as race/ethnicity and free/reduced price lunch status. It is also possible that some schools are worse at reporting data than others. Large schools, for example, may be less diligent about data entry due to the larger scale of the task. For these reasons, dropping all cases that have missing data is likely to create a sample of students that is not representative of the overall New York City public school population, biasing my results. Finally, because a substantial portion of students are missing data on just some of their measures, dropping these students means losing information on the measures for which they do have data as well. The result would be a loss of potentially valuable information (Allison, 2002; Rubin, 1987). Prior to imputation, rates of missing data in my analytic sample ranged across measures. Some measures were missing no data or minimal data, for example less than one percent of cases were missing data on race/ethnicity or gender. However, 12.8 percent of cases were
missing data on whether students were overage for eighth grade and a high of 14.2 percent of cases were missing data on mobility during middle school.

**Multiple imputation.** I use multiple imputation to address missing data in my dataset, using the MI package in STATA to generate five imputed datasets. 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). Multiple imputation has been shown to perform better than other common methods of addressing missing data both generally (Rubin, 1996), and in the specific context of propensity score matching methods (Mattei, 2009). I impute separately for the 2005-06 and the 2006-7 cohorts to allow for changes in the associations between student characteristics over time. I conduct matching and estimate my results separately on each of five imputed datasets, and then combine the results according to Rubin’s rule (Rubin, 1987). In my imputation models I include all variables that are used in later analyses.

**Drop-outs.** Some data are missing for known reasons. Thanks to NYC DOE discharge data, we have indicators of why students leave schools, specifically whether they drop out. These students pose problems for analyses on school effects because they do not receive the full four-year “treatment” of attending high school. How one deals with drop-outs is more than a technical concern—it fundamentally affects the interpretation of the study’s results. For example, one strategy would be to focus on the progression through the math pipeline only for students who remained at the same school for four years. This would be informative in terms of evaluating math course-taking patterns among *graduates* of small versus large schools, but it would reduce the sample, eliminating the most disadvantaged and academically underprepared students. Furthermore, since small schools tend to have lower drop-out rates than large schools (see Bloom, et al., 2010), and students who drop out tend to fall lower on the achievement distribution, small schools would be unfairly
penalized by a strategy that evaluates only the outcomes of students who remain in school for a full four years.

Another strategy would be to assess the most advanced level of math that students reach regardless of whether they drop out or not. Following this method, 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 remain in school. The downside is that we do not know how far these students would have made it through the curriculum had they remained in school. Nonetheless, I choose the latter strategy, because I believe that excluding drop-outs would unfairly advantage schools with low retention rates, when in reality, this is the worst educational outcome of all. I test the sensitivity of these results by re-analyzing my analytic, propensity score matching results on a sample of students that excludes drop-outs (presented in Chapter 6). Drop-outs make up four percent of my analytic sample. See Figures A.1 and A.2 in the Appendix for the distribution of the math pipeline measure with and without drop-outs included in the sample. I designated students as having dropped out if their administrative discharge code indicated that they dropped out or if they were missing all school data for consecutive semesters and did not return to high school in my sample window.

**Descriptive Approach**

My first research question focuses on describing the curricular structures in New York City high schools, focusing on whether the new, small schools implement different curricular structures than the alternate types of public high schools in New York City. Students cannot take courses that are not offered, so describing course offerings is the first step in assessing differences in educational opportunities. This is also a key step in evaluating whether small schools actually implement a narrower, more college preparatory curriculum than their larger counterparts. The descriptive analyses of schools’ student body compositions and curricular structures are the most encompassing
analyses in this study with regards to the types of schools included. I will examine the student body compositions and curricular structures of the following categorizations of high schools:

- Specialized high schools
- Large schools
  - Academically selective
  - Academically nonselective
- Midsize schools
  - Academically selective
  - Academically nonselective
- Small schools
  - Academically selective
  - Academically nonselective
    - New, small, nonselective schools
    - Veteran, small, nonselective schools

The next step in my descriptive analyses is to explore observed differences among students, differentiated by groups that characterize their progress through the math pipeline. I compare students’ socio-demographic and academic achievement characteristics for three groups: (1) “low”: those whose most advanced math course passed was below Algebra II/Trigonometry in difficulty [32.9 percent], (2) “middle”: those who completed both semesters of Algebra II/Trigonometry but did not reach Pre-Calculus [44.5 percent], and (3) “high”: those who reached Pre-Calculus or beyond [22.6 percent].

Similar to exploring observed differences in groups of students differentiated by their most advanced math course, it is also helpful to examine the differences in school characteristics exhibited by groups of schools differentiated by the average progress of their students through the math curriculum. Thus, I compare school demographics and curriculum structure for three groups of schools: (1) those with low-average math course-taking (more than one standard deviation below the
mean), (2) those with middle-average course-taking (within one standard deviation of the mean), and (3) those with high-average course-taking (more than one standard deviation above the mean).

**Multilevel Modeling.** Since my second research question focuses on the relationship between school size and student outcomes—essentially, a “school effects” study—I use multilevel modeling, which was designed for exactly this type of nested data structure. Multilevel modeling has been used in other studies seeking to address the relationships between structural features of schools and students’ course-taking patterns (see for example: Lee et al, 1998; Lee, Croninger, & Smith, 1997).

To explore research question 2 regarding students’ progress through the math pipeline, I use two-level, multivariate models.

The first step in multilevel modeling is to partition the variance in the outcome into its within- and between-school components. It is only the portion of the total variance that lies between schools that can be modeled as a function of school factors such as size. Thus, before beginning my analyses for research question 2, I compute the intraclass correlation (ICC) for my outcome, progress through the math pipeline. The ICC represents the proportion of the variability in a student-level outcome that is systematically between schools. I compute the ICC from a fully unconditional multilevel model, meaning that no predictors are specified at either the student- or school-level. Second, I construct a student-level model, which establishes the overall relationships between students’ characteristics and their progress through the mathematics pipeline, independent of school context (although the nested nature of the data is accounted for). Third, I conduct a between-school multilevel model, which explores the relationships between school characteristics such as curriculum structure and student body composition, and students’ progress through the math curriculum. In this model, I disentangle the associations between school size, selectivity, and era of founding from one another. And finally, in my most complex, between-school models, I directly estimate the effect of attending a new, small, nonselective school on students’ progress.
through the math pipeline and explore whether the effect of attending a NSN school varies by students’ cohort, race/ethnicity, or entering math achievement. To take one of the most complex models as an example, the model for entering math achievement can be written as:

\[
MTHPIPE_{ijh} = \beta_{j} + \beta_{Mh}\left(MATH8th_{ih} - MATH8th_{h}\right) + \beta_{Nh}\left(ASIAN_{ih} - ASIAN_{h}\right) + \\
\beta_{Bh}\left(BLACK_{ih} - BLACK_{h}\right) + \beta_{Hh}\left(HISPANIC_{ih} - HISPANIC_{h}\right) + \\
\beta_{NH}\left(NATIVE / MULTI_{ih} - NATIVE / MULTI_{h}\right) + \\
\beta_{Coh}\left(COHORT0607_{ih} - COHORT0607_{h}\right) + \beta_{X_{ih}X_{h}} + \ldots + \beta_{X_{ih}X_{h}} + r_{ih}
\]

Level 1:

\[
\beta_{0h} = \gamma_{00} + \gamma_{01}\left(NSN_{h}\right) + u_{0h}
\]

\[
\beta_{Mh} = \gamma_{M0} + \gamma_{M1}\left(NSN_{h}\right) + u_{Mh}
\]

\[
\beta_{Nh} = \gamma_{N0} + u_{Nh}
\]

\[
\beta_{Bh} = \gamma_{B0} + u_{Bh}
\]

\[
\beta_{Hh} = \gamma_{H0} + u_{Hh}
\]

\[
\beta_{NH} = \gamma_{NH0} + u_{NHh}
\]

\[
\beta_{Coh} = \gamma_{C0} + u_{Coh}
\]

The Level-1 model estimates progression through the math pipeline for student \(i\) in high school \(h\) as a function of their entering eighth grade mathematics achievement level and socio-demographic characteristics. Random slopes are included for entering math achievement, racial/ethnic groups, and the 2006-07 cohort. These decisions were made because extant literature suggests that attending a small school might be more beneficial for low-achieving, low-income, and minority students. The random slope for the 2006-07 cohort was included to allow for the fact that the effect of attending a NSN school might change the longer the policy is in place. A random slope for students eligible for free or reduced price lunch was also tested but found to be non-significant. At the school level, students’ progress through the mathematics pipeline is modeled as a function of attending a NSN school, which takes on the value of 1 for students who attend a NSN school; \(\mu_{0h}\) represents the error associated with high school \(h\). I investigate the effectiveness of new, small, nonselective schools by estimating the effects on the intercepts. Finally, I examine whether the effect
of attending a small school varies by initial math achievement level. Specifically, I model the entering math achievement effect (the student-level slope) on my outcome as a function of attending a new, small, nonselective school (the slopes-as-outcomes approach commonly used in multilevel analyses; Seltzer, 1995). Across the multilevel descriptive models, I group-mean center race/ethnicity indicators and students’ entering mathematics achievement, meaning that these estimates specifically compare students within the same schools. All other student-level measures are grand-mean centered. At the school level, dummy-coded measures are left un-centered and continuous measures are grand-mean centered.

**Analytic Approach**

**Causal Framework.** In addition to examining the relationships between student and school characteristics and students’ math course-taking, I will also assess the impact of small schools on students’ progression through the math curriculum (research question 2). The first step in estimating a causal impact is clarifying the treatment group. In this study, the treatment of interest is attending a new, small, nonselective (NSN) school. The next step in estimating causal effects is identifying the counterfactual—that is, what would have happened in the absence of the treatment. However, I argue that in the case of small school reform in New York City, there are two plausible counterfactuals: the contemporary counterfactual, which assumes that if small high schools didn’t exist, then the students who currently attend them *would have* attended another available, mid-sized or large school, or; the historical counterfactual, which assumes that if the entire small school reform policy had never been implemented, and the large, “failing” schools had never been replaced by small schools, then students currently attending new small schools (or at least some of them) *would have* attended the large “failing” high schools. This study assesses the impact of NSN schools relative to the contemporary counterfactual. I assume that if small school students were not able to
matriculate into their small, nonselective school in ninth grade, they would have attended another high school open that year.

Two immediate concerns arise. First, over the decade of small school reform in New York City, new small schools were opening as large, failing high schools were closing. Thus neither the treatment group (NSN schools) nor the control group (non-NSN schools) constituted a consistent group. Many of the worst performing large schools were closed, and even the large schools that existed throughout the entire period were likely changing. The “effect” of attending a small school, relative to the contemporary alternatives, could potentially be different for different cohorts. To address this issue, I conduct my analyses separately for each cohort.

The second challenge is the potential for selection bias. The concern is that students who attend small schools might differ from their counterparts who attend large schools on several unmeasured dimensions, such as motivation or parental support, that independently affect progression through the math curriculum. Put simply, do small schools cause students to take more advanced courses or would the types of students who attend small schools take more advanced coursework no matter what size school they attended? To address selection bias, I utilize several propensity score matching techniques.

*Propensity score matching.* Propensity score analyses involve matching cases on their probability of receiving the treatment—in this case attending a small, non-selective high school—based on all other observed information for that case prior to the treatment (Gelman & Hill, 2007; Rubin & Thomas, 1996; Rubin, 1997). Essentially, the predicted probabilities—the propensity scores—measure the characteristics of students who attend small schools, permitting the identification of similar students in large schools. The propensity for attending a small school is estimated through a

---

11 Hemphill & Nauer (2009) find that the closure of the initial round of large high schools resulted in increased enrollments at the remaining large schools.
logistic regression, where the outcome is the probability of attending a small school as a function of all measured, pre-treatment variables that could potentially cause selection effects (the confounding covariates). Propensity score matching is an improvement over traditional linear regression techniques, which control for confounding covariates, because it creates a more balanced sample of treatment and control cases, a method that does not require the model to extrapolate over areas with minimal data (Gelman & Hill, 2007; Rosenbaum & Rubin, 1983).

I conduct three separate propensity score techniques to ensure that my estimates are robust across model specifications. First I generate propensity scores through a multilevel, logistic regression that utilizes both student characteristics and the attributes of their middle schools to predict students’ propensity to attend a NSN school. The model can be written:

\[
\text{Level 1: } \logit(\text{NSN}_{im}) = \beta_{Mn}(MATH^{8th}) + \beta_{im}X_{im} + \ldots + \beta_{kn}X_{kn} + r_{im}
\]

\[
\text{Level 2: } \beta_{im} = \gamma_{00} + \gamma_{01}W_m + \ldots + \gamma_{0k}W_m + u_{im}
\]

Here \(\text{NSN}_{im}\) is the probability of attending a new, small, nonselective high school for student \(i\) in middle school \(m\) and is modeled as a function of both student-level covariates including their eighth grade math achievement (∙\(MATH^{8th}\)) and socio-demographic and behavioral characteristics (∙\(X_1\) through \(X_k\)) as well as middle school characteristics (the \(W_m\)s) such as student enrollment, borough, Title 1 status, aggregate student characteristics, and aggregate teacher characteristics. Utilizing a multilevel model adjusts the propensity score for school as well as student characteristics.

To assess whether the propensity score was properly specified, I utilize balance diagnostics to reveal whether the means and distributions of measured baseline covariates are similar between treated and control groups. To take into account both means and distributions of the covariates, I compare the standardized difference in means—the means divided by the standard deviation of the treatment group—between the matched samples (Stuart, 2010). Research suggests that sufficient
balance can be reached with an absolute standardized difference in means of no more .25 between the matched samples (Rubin, 2001; Stuart, 2010). Others suggest a more stringent threshold of a .1 standardized difference in means (Austin, 2011; Normand et al., 2001). My matched samples do not exceed an absolute standardized difference of .05 on any measure. I also ensure that the ratio of treatment and control group standard deviations is between .90 and 1.1 for all continuous variables and that dummy variables do not differ by more than .025 percent between groups. Figures A.3 and A.4 in Appendix A display the absolute standardized differences by student characteristic for each of the five imputed datasets by cohort.

Once sufficient balance was achieved, I use these propensity scores in two different ways. First, I use them to select a matched group of control cases from the group of students who do not attend a NSN school to serve as a comparison group for the students in NSN schools. I use one-to-one, nearest neighbor matching, without replacement and I restrict analyses to areas of common support. Nearest-neighbor matching performs particularly well with datasets that include many more control units than treatment units, and if many of the controls are quite different than the treated individuals, which is the case with these data (Ho et al., 2007; Stuart, 2010). Because I discarded unmatched control students, my estimates are the effect of the treatment on the treated. As a robustness check I also test matching with replacement and find comparable results. I then use further regression adjustment, controlling for all pre-treatment, student-level characteristics within this matched sample. Since propensity score matching generally produces very similar, but not identical, treatment and control groups, analyzing the matched samples using regression models with additional controls helps minimize any bias due to inexact matching. I refer to this method as multilevel propensity score matching (PSM).

Next, I use the propensity scores as modeling weights in a regression that includes all students in my analytic sample. Essentially, I weight the control units in such a way that they reflect
the same propensity to attend a NSN school as the treatment group. I refer to this strategy as multilevel propensity score weighting (PSW). Utilizing propensity scores as model weights is a more efficient variation on matching that does not require unmatchable cases to be excluded (Hirano, Imbens & Ridder, 2003; Morgan & Harding, 2006; Crosnoe, 2009). With this strategy, standard regression models (in this case, multilevel models) are weighted by the propensity scores, which over- or undercounts control cases in the sample based on their propensity to be in the treatment group. The equation for transforming a propensity score into a modeling weight is: \( w(t, x) = t + (1 - \tilde{t})/[1 - e(x)] \), where \( e(x) \) is the propensity score, \( t \) is treatment status and \( x \) represents the covariates (Hirano and Imbens, 2001). The resulting weight captures the treatment effect on the treated.

Finally, in my preferred approach, I regenerate students’ propensity scores using only student characteristics, and then force an exact match on students’ middle schools when selecting my matched sample of controls. I utilize the same balance diagnostics and again, my matched samples do not exceed an absolute standardized difference of .05 on any measure and meet all other requirements discussed above. Figures A.5 and A.6 in Appendix A display the absolute standardized differences by student characteristic for each of the five imputed datasets by cohort. As before, I use one-to-one, nearest neighbor matching, without replacement and I restrict analyses to areas of common support. I then perform additional regression adjustment within this matched sample. This technique restricts comparisons to students who attended eighth grade at the same middle school in the same year but subsequently attended different sized high schools. This strategy removes the threat of bias from unmeasured covariates that are related to students’ middle school experience, and reduces the threat of bias from neighborhood effects, since students tend to attend middle school near their home.

Asumptions for causal interpretation. Both my multilevel descriptive analyses and my propensity score matching approach assume that there is no omitted variable bias and that the Stable Unit
Treatment Value Assumption (SUTVA) holds. The assumption of no omitted variable bias is more believable in the propensity score models than in the descriptive models, particularly the technique that forces an exact match on students’ middle schools. By matching on students’ middle schools I have eliminated much of the bias associated with students’ middle-school educational experiences. Nonetheless, the assumption of no omitted variable bias is still a very strong assumption given the level of student selection into high schools in New York City and the limited number of available student-level controls. The SUTVA assumption, which requires that the potential outcome of one student should be unaffected by the assignment of treatments to the other students, is also a strong supposition in this situation, given the well-documented phenomenon of peer-effects in secondary school research.

In the multilevel descriptive analyses, there is a further assumption that my parametric model specification is correct. Propensity-score matching relaxes this assumption, but adds another (less heroic) assumption of its own: that there is sufficient balance and overlap between the treatment and control groups to support my model extrapolations. Given the tenability of the assumptions discussed above, I do not interpret the results from my multilevel descriptive analyses as causal and I interpret my propensity score models as causal very cautiously.

**Overview of Findings Chapters**

In the following chapters, I present the descriptive and analytic results generated from the analyses outlined above. In Chapter 5, I situate the new, small, nonselective schools within the broader context of New York City public high schools by comparing their student body compositions and curricular structures across schools. I also explore observed differences among students, differentiated by groups that characterize their progress through the math pipeline. Next, I examine the differences in school characteristics exhibited by groups of schools differentiated by the average progress of their students through the math curriculum.
In Chapter 6, I first present my multilevel descriptive results, which examine the student and school characteristics associated with students’ success in the math curriculum. I also investigate whether the effect of attending a NSN school on students’ progress through the math curriculum may depend on their background characteristics such as race, class, or entering achievement. Next, I present my analytic results from the propensity score matching analyses, which assess whether attending a NSN school improved students’ progress through the mathematics curriculum relative to the other school options available in New York City. Finally, I go beyond course titles to investigate how the rigor of the curriculum in NSN schools compares to the alternative New York City public high school options.
Chapter 5
Small Schools in a Stratified System

“This study shows conclusively that our new small high schools changed thousands of lives in New York City, across every race, gender and ethnicity — not only helping them graduate, but graduate ready for college. When we see a strategy with this kind of success, we owe it to our families to continue pursuing it aggressively.”

~Chancellor Dennis M. Walcott
January 26, 2012

In the above quote the Chancellor of New York City Public Schools, citing the MDRC study (Bloom et al., 2010; 2012) that found that the newly created small high schools improved graduation rates as well as some measures of college readiness, portrays the city’s small high school reform efforts as an unparalleled success and vows to continue implementing it “aggressively.” The results that I present in the next two chapters reveal that the real story is both more nuanced and less rosy than Chancellor Walcott suggests. Proponents of small school reform promised that these schools, which targeted underperforming students living in the most disadvantaged areas of the five boroughs, would provide their students with a more rigorous curriculum and would produce more college and career-ready graduates. A decade later, the Chancellor of schools proclaimed that these promises had been fulfilled based on evidence from one study. As I discussed in more detail in my literature review in Chapter 2, this MDRC study (Bloom et al., 2010; 2012) painted a much more complex and mixed picture. While the researchers did find that the new small high schools improved graduation rates as well as some degrees of college readiness—as measured by Regents scores in English—absent from the Chancellor’s speech was the fact that the study found no improvement in students’ college readiness in Mathematics. As small school reform continues to expand in New York City and is increasingly held up as an urban education strategy worthy of emulation, it is crucial for policymakers to understand why these schools are failing to improve students’ math performance.
In the next two chapters, I delve deeper into the role that the new small schools play in their students’ access to an advanced, college preparatory math curriculum and their success in taking and passing more advanced math courses. In doing so, my study builds on the existing findings on small schools in New York City to assess the impact of small schools on students’ college readiness from a different perspective—student course-taking. My results shed light into the “black box” of within-school mechanisms that explain why the reform may not be improving disadvantaged students’ college readiness in mathematics. My first step in evaluating the success of the reform is to describe the types of students who attend the new small schools and situate them within the context of the broader landscape of public high school options in New York City. Clearly, in order for small schools to improve disadvantaged students’ academic outcomes, they must first enroll disadvantaged students.

After detailing the differences in schools’ student body composition, I will explore the curriculum structures that are put in place by different types of high schools. The constrained curriculum hypothesis holds that if small schools implement a narrower, more academic curriculum, their students (particularly disadvantaged students) will take more advanced courses than they would in a school with a more differentiated curriculum. Thus, in Chapter 6, I assess this assumption in the context of New York City small school reform. First I employ multivariate, multilevel models to reveal which student and school characteristics are associated with students’ success in the math curriculum. I also investigate whether low-income, minority, and low-achieving students are differentially impacted by attending a new, small school, as the literature suggests. Next, I present the results of my propensity score matching analyses, which assess whether attending a new, small, nonselective school improved students’ progress through the mathematics curriculum relative to the other school options available in New York City. Finally, I go beyond course titles to assess whether
the rigor of the curriculum varies across high schools in New York City as a function of characteristics such as size, selectivity, and student body composition.

**Descriptive Results**

**The Geography of Small School Creation**

As previewed in the literature review, my results confirm that small school reform was not implemented evenly throughout New York City. Instead, the creation of small schools has been heavily concentrated in lower-income areas of the city, particularly in the Bronx and central Brooklyn (see Table 5.1). This is no surprise given that the architects of the reform conceptualized the new small schools as replacements for the large, under-performing high schools, which were selected for closure due to their low graduation rates and were largely located in areas with median incomes of less than $40,000 (Quint et al., 2010).

As presented in Table 5.1, the freshman cohort entering high school in the fall of 2005 (on track to graduate in 2009) confronted a wide array of high school options, nearly one third of which were newly created small, nonselective schools. By 2005-06, only 14 small, nonselective schools remained that had been created prior to the 2002-2003 school year, while 103 had been established since. Indeed NSN schools represent over 70 percent of all new schools created during the three-year window between the fall of 2002 and the fall of 2005 in which the policy had been implemented. No large schools were created during this time period.

Furthermore, the new, small schools were intentionally created to be academically nonselective because they were designed to serve the same neighborhoods and student populations as the large, nonselective schools, which they replaced. This also explains why three-quarters of NSN schools are located either in Brooklyn or the Bronx. It is also important to note that, as small schools continue to replace large schools in New York City, the landscape of alternative options
(non-NSN high schools) is also shifting. Clearly, as large nonselective schools are shuttered and no new ones are created, there are fewer large high school options available to incoming freshmen. Moreover, a substantial portion of the large nonselective schools that do remain are located in less disadvantaged, whiter areas of the city than their closed counterparts. For example, nearly half of all large nonselective schools are situated in either Queens or Staten Island, while less than eight percent of small, nonselective schools are located in those two boroughs. In addition, nearly half of all large selective schools are sited in Manhattan. From an evaluation perspective in which NSN schools are conceptualized as the “treatment” and the alternative high school options are conceptualized as the “control,” this means that neither treatment nor control groups are constant across cohorts. Indeed, as the worst performing large high schools are closed, the remaining population of large schools are higher performing, on average, than the original population, prior to the reform. Thus, my analyses estimating the effect of attending a NSN school on student course-taking relative to contemporary, alternative high school options will be a conservative estimate of the true effect of the reform, as the worst of the large high schools are no longer in existence.

Turning now to the 2006-07 entering freshman cohort (class of 2010), the overall landscape of high school options is similar to the year before. However, 20 new schools opened for the 06-07 school year: 13 NSN schools, six academically selective small schools, and one specialized high school. In addition, one large, nonselective school (Adlai Stevenson High School) and one small selective school (The New School for Arts and Sciences), both located in the Bronx, began phasing out, no longer accepting entering freshmen.

For both cohorts, midsize schools are predominantly academically selective (only four nonselective midsize schools were operating in either year). Because this group is so small, it is omitted from subsequent descriptive tables comparing schools by size and selectivity. However, these schools do remain in the sample.
Table 5.1: School Location and Creation Era, Differentiated by Size, Selectivity, and Freshman cohort

**2005-2006**

<table>
<thead>
<tr>
<th>Borough</th>
<th>New, Small, Nonselective</th>
<th>Veteran, Small, Nonselective</th>
<th>Small, Selective</th>
<th>Midsize, Nonselective</th>
<th>Midsize, Selective</th>
<th>Large, Nonselective</th>
<th>Large, Selective</th>
<th>Specialized</th>
<th>All Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronx</td>
<td>52</td>
<td>5</td>
<td>20</td>
<td>0</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>95</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>27</td>
<td>3</td>
<td>19</td>
<td>0</td>
<td>16</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td>86</td>
</tr>
<tr>
<td>Manhattan</td>
<td>16</td>
<td>5</td>
<td>33</td>
<td>2</td>
<td>12</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>83</td>
</tr>
<tr>
<td>Queens</td>
<td>7</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>8</td>
<td>15</td>
<td>4</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>Staten Island</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School opened</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before 02-03</td>
<td>0</td>
<td>14</td>
<td>51</td>
<td>4</td>
<td>39</td>
<td>44</td>
<td>21</td>
<td>5</td>
<td>178</td>
</tr>
<tr>
<td>During or After 02-03</td>
<td>103</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>144</td>
</tr>
</tbody>
</table>

**2006-2007 (Change +/-)**

<table>
<thead>
<tr>
<th>Borough</th>
<th>New, Small, Nonselective</th>
<th>Veteran, Small, Nonselective</th>
<th>Small, Selective</th>
<th>Midsize, Nonselective</th>
<th>Midsize, Selective</th>
<th>Large, Nonselective</th>
<th>Large, Selective</th>
<th>Specialized</th>
<th>All Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronx</td>
<td>+4</td>
<td>-</td>
<td>-1</td>
<td>-</td>
<td>-</td>
<td>-1</td>
<td>-</td>
<td>-</td>
<td>+2</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>+6</td>
<td>-</td>
<td>+2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+1</td>
<td>+9</td>
</tr>
<tr>
<td>Manhattan</td>
<td>+1</td>
<td>-</td>
<td>+2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+2</td>
</tr>
<tr>
<td>Queens</td>
<td>+2</td>
<td>-</td>
<td>+2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+5</td>
</tr>
<tr>
<td>Staten Island</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

| Net Change          |                           |                              |                  |                       |                    |                    |                  |             |             |
| from 05-06          | +13                       | -                            | +5               | -                     | -1                 | -                  | +1               | +18         |

a See Measures section for descriptions of how school size and selectivity categories are calculated.
Who Goes Where?

If the principle objective of small high school reform in New York City is to improve the academic experience and achievement of disadvantaged, under-achieving students, then a necessary intermediary goal is to create small schools which enroll this type of student. In this regard, the Department of Education (DOE) was successful. Table 6.2 reports the socio-demographic and achievement-based characteristics of first-time ninth-grade students enrolled in New York City public high schools (differentiated by size and selectivity) for the 2005-06 cohort. The distributions of socio-demographic and achievement-based characteristics across school types remain largely consistent between the two cohorts. Thus, to avoid repetition, for the remainder of the descriptive statistics in this chapter I present only the tables for the 2005-06 cohort in the text. The equivalent tables for the 2006-07 cohort can be found in Appendix A. I will refer to these tables only to reference the elements that do change between cohorts.

New small schools serve most disadvantaged students. Beginning with the 2005-06 cohort (See Table 5.2), just over 90 percent of students attending a NSN school are black or Hispanic, although these two groups make up under three-quarters of all NYC students. On the other end of the spectrum, black and Hispanic students make up only 19 percent of the specialized high school population, which is disproportionately Asian (half) and white (one third). Interestingly, black and Hispanic students are also underrepresented in both large selective and large nonselective high schools, making up under two-thirds of the student population in either category. Indeed, large nonselective schools have a lower black and Hispanic student population than do small selective schools (64.2% versus 81.5%). As discussed earlier, this is likely due to the fact that many of the large nonselective schools that used to be in low-income, high-minority areas of the Bronx and central Brooklyn have been closed, leaving a relatively more affluent and whiter group of large schools. It is also interesting to note that, while veteran, small, nonselective (VSN) schools also serve high
proportions of minority students, they have a somewhat higher percentage of Hispanic students and a lower percentage of black students than the NSN schools.

The proportions of low-income students eligible for free or reduced price lunch follow a similar pattern to race/ethnicity, with small nonselective schools (both new and veteran) having the highest proportion, followed next by small and midsize selective schools, then by large selective and nonselective schools, and finally by specialized high schools. Beyond racial and economic composition, NSN schools also serve higher than average proportions of students who are considered overage for eighth grade and who had higher rates of absences and tardies in eighth grade. Both NSN and large nonselective schools enroll relatively high proportions of overage students (just over 20% in both cases) compared to small, midsize, and large selective schools (all hovering around 15%) and specialized high schools (only 2.5%). Veteran, small, nonselective schools enroll the highest proportion of overage students (nearly one-third). The pattern for eighth grade absences is somewhat similar, with students in NSN and large nonselective schools averaging roughly 17 absences, compared to roughly 15 for VSN schools as well as small and midsize selective schools, 13 for large selective, and 7 for specialized high schools. Students in NSN schools also registered the highest number of tardies, on average, in eighth grade.

**Small, nonselective schools at bottom of academically stratified system.** Another important characteristic of schools’ student body compositions is the level of academic achievement that students possess upon entering 9th grade. As presented in Table 5.2, NSN schools enroll students whose eighth grade math and English Language Arts (ELA) skills are roughly a quarter of a standard deviation below average. Unsurprisingly, entering achievement appears to be correlated most with the selectivity of high schools. The 14 remaining veteran, small, nonselective schools have the lowest entering achievement of any school category, enrolling students who are over two-fifths of a standard deviation below average in both math and ELA achievement. Large nonselective
schools also enroll freshmen with below-average achievement (although not as low as new or veteran small nonselective schools). Small and midsize schools enroll students who are just slightly above average in both math and reading achievement, while large selective schools enroll students who are, on average, almost half a standard deviation above the mean. Finally, specialized high schools enroll students who are a staggering 1.62 standard deviations above the mean, on average.

The most notable difference between the 2005-06 and 2006-07 cohorts is that the eighth grade achievement levels of the freshmen entering NSN schools fall farther below average in 06-07 than in the previous year (see Table A.7 for the 06-07 cohort). In 05-06 NSN students began high school with math achievement, on average, just over a quarter of a standard deviation below the mean (-0.277), however in 06-07 they began almost two-fifths (-0.384) below. The story is the same for their ELA achievement: -0.229 standard deviations in 05-06 and -0.295 standard deviations in 06-07. Test scores are standardized within cohort so these figures do not speak to students’ absolute level of content mastery, but rather their level of mastery relative to the rest of the students in their cohort. This drop in relative achievement highlights the importance of accounting for students’ cohort in future multivariate analyses.

**Evidence of policy accommodations fading.** Two categories in which NSN schools do not record the highest proportions are special education students and English Language Learners (ELL). Given the history of small school reform in New York City, this is not surprising. One of the accommodations granted to new small schools during their initial two start-up years was that they were not required to enroll ELL or special education students. As displayed in Table 5.2, by the freshmen cohort of 2005-06, NSN schools’ ELL and special education enrollments were proportionate to the system-wide averages, although still lower than large nonselective schools with regards to the proportion of special education and ELL students and lower than all other nonselective schools with regards to ELL students. Indeed, the most substantial difference between
NSN and VSN schools is the proportion of ELL students, with VSN schools serving approximately three times the number of ELL students than their newer counterparts. However, by the 2006-07 cohort the proportion of special education and ELL students attending NSN schools rises above the system-wide average (although the proportion of ELLs is still substantially lower than in VSN schools). This is likely due to more of the new small schools exiting their initial grace period of exemptions from enrolling these categories of students.

**Summary and implications.** To review, I find that NSN schools did enroll a substantially more disadvantaged group of students than would be the case if students were sorted randomly across school types in New York City. NSN schools serve higher proportions of low-income, low-achieving, overage, and black and Hispanic students than the system as a whole. They also enroll students with more absences and tardies in eighth grade and by 06-07 they enroll higher percentages of special education and ELL students. Moreover, it is clear that these results for NSN schools are symptomatic of system-wide patterns of stratification related to schools’ size, selectivity, and geographic location. In some cases the differences in student demographics are incredibly severe. For example, the students matriculating to specialized high schools in 06-07 began ninth grade with math skills almost two full standard deviations more advanced than students matriculating to NSN schools. This high degree of stratification is not surprising, given the way that high school assignment is structured in New York City. A wide range of high school options coupled with an institutionalized policy of both students choosing schools and schools choosing students creates prime conditions for the sorting of students across schools. The findings presented in Tables 5.2 and A.7 are evidence of the sorting process and underscore the need for any evaluation of school effects in New York City to account for the high levels of student selection into schools.

I also find that, while the new, small, nonselective schools are more disadvantaged than most other school categories, they are comparable in this regard to the few remaining veteran, small,
nonselective schools. In some key categories such as entering achievement and the proportion of ELL students, these VSN schools are substantially worse off. However, the disparities that exist between NSN and VSN schools in the earlier cohort appear to have equalized somewhat by the later cohort, suggesting that the student body of NSN schools is, if anything, becoming even more disadvantaged over time.

It is also important to note that while serving low-income, minority, and under-achieving students was a goal of the DOE’s small high schools initiative, the literature on student body composition suggests that racially isolated schools, schools of concentrated poverty, and schools with predominantly low-achieving students can all have distinctly negative effects on individual students’ academic outcomes (Ready & Silander, 2011). Furthermore, as discussed in my conceptual framework, I hypothesize that student body characteristics, particularly average entering achievement will mediate the relationship between school size and curricular structure, complicating the constrained curriculum hypothesis, which assumes a direct link from school size to curricular structure to student achievement. Thus going forward, it will be critical to my analyses to monitor the role that student body composition plays in schools’ curricular structures and students’ progress through the math curriculum.
Table 5.2: Socio-Demographic and Achievement-Based Characteristics of First-Time Ninth-Grade Students Enrolled in DOE High Schools by Size and Selectivity, 2005-2006 School Year

<table>
<thead>
<tr>
<th></th>
<th>New, Small, Nonselective&lt;sup&gt;a&lt;/sup&gt; (N=103)</th>
<th>Veteran, Small, Nonselective&lt;sup&gt;a&lt;/sup&gt; (N=14)</th>
<th>Small, Selective&lt;sup&gt;a&lt;/sup&gt; (N=45)</th>
<th>Midsize, Selective&lt;sup&gt;a&lt;/sup&gt; (N=44)</th>
<th>Large, Nonselective&lt;sup&gt;a&lt;/sup&gt; (N=44)</th>
<th>Large, Selective&lt;sup&gt;a&lt;/sup&gt; (N=25)</th>
<th>Specialized&lt;sup&gt;b&lt;/sup&gt; (N=8)</th>
<th>All Schools&lt;sup&gt;b&lt;/sup&gt; (N=322)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48.1</td>
<td>47.6</td>
<td>43.7</td>
<td>51.5</td>
<td>50.9</td>
<td>50.3</td>
<td>49.9</td>
<td>49.7</td>
</tr>
<tr>
<td>Female</td>
<td>51.9</td>
<td>52.4</td>
<td>56.3</td>
<td>48.5</td>
<td>49.1</td>
<td>49.7</td>
<td>50.1</td>
<td>50.3</td>
</tr>
<tr>
<td>Race/Ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>4.10</td>
<td>7.85</td>
<td>8.63</td>
<td>8.22</td>
<td>17.0</td>
<td>24.1</td>
<td>48.6</td>
<td>14.4</td>
</tr>
<tr>
<td>Black</td>
<td>40.5</td>
<td>26.0</td>
<td>40.4</td>
<td>34.2</td>
<td>29.4</td>
<td>25.8</td>
<td>9.64</td>
<td>32.0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>51.0</td>
<td>61.1</td>
<td>41.1</td>
<td>45.7</td>
<td>34.8</td>
<td>37.0</td>
<td>9.15</td>
<td>40.0</td>
</tr>
<tr>
<td>Native/Multi</td>
<td>0.64</td>
<td>0.63</td>
<td>0.66</td>
<td>0.58</td>
<td>0.67</td>
<td>0.55</td>
<td>0.54</td>
<td>0.63</td>
</tr>
<tr>
<td>White</td>
<td>3.60</td>
<td>4.16</td>
<td>8.31</td>
<td>10.6</td>
<td>18.1</td>
<td>12.4</td>
<td>31.8</td>
<td>12.8</td>
</tr>
<tr>
<td>Eligible for free or reduced-price lunch (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>86.2</td>
<td>89.1</td>
<td>78.5</td>
<td>78.4</td>
<td>74.0</td>
<td>75.7</td>
<td>52.3</td>
<td>77.3</td>
</tr>
<tr>
<td>Overage for 8th grade (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21.4</td>
<td>31.1</td>
<td>15.1</td>
<td>14.9</td>
<td>20.4</td>
<td>13.9</td>
<td>2.54</td>
<td>18.4</td>
</tr>
<tr>
<td>Special education (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.3</td>
<td>11.3</td>
<td>10.3</td>
<td>10.1</td>
<td>12.5</td>
<td>8.08</td>
<td>0.80</td>
<td>10.8</td>
</tr>
<tr>
<td>English Language Learner (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.7</td>
<td>42.5</td>
<td>7.6</td>
<td>9.10</td>
<td>16.0</td>
<td>9.12</td>
<td>1.76</td>
<td>13.2</td>
</tr>
<tr>
<td>Days absent in 8th grade (Mean)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.2</td>
<td>15.8</td>
<td>14.3</td>
<td>14.9</td>
<td>17.8</td>
<td>12.5</td>
<td>7.20</td>
<td>15.8</td>
</tr>
<tr>
<td>Days tardy in 8th grade (Mean)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.6</td>
<td>18.2</td>
<td>16.6</td>
<td>15.3</td>
<td>16.9</td>
<td>12.4</td>
<td>5.30</td>
<td>16.1</td>
</tr>
<tr>
<td>Entering Math achievement&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.277</td>
<td>-0.431</td>
<td>0.039</td>
<td>0.077</td>
<td>-0.155</td>
<td>0.458</td>
<td>1.58</td>
<td>0.00</td>
</tr>
<tr>
<td>SD</td>
<td>0.864</td>
<td>0.878</td>
<td>0.879</td>
<td>0.939</td>
<td>0.966</td>
<td>1.10</td>
<td>0.698</td>
<td>1.00</td>
</tr>
<tr>
<td>Entering ELA achievement&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.229</td>
<td>-0.444</td>
<td>0.066</td>
<td>0.076</td>
<td>-0.173</td>
<td>0.413</td>
<td>1.62</td>
<td>0.00</td>
</tr>
<tr>
<td>SD</td>
<td>0.782</td>
<td>0.826</td>
<td>0.926</td>
<td>1.01</td>
<td>0.911</td>
<td>1.17</td>
<td>0.998</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<sup>a</sup>See Measures section for descriptions of school size and selectivity classifications.

<sup>b</sup>Column school counts do not sum to 322 because Specialized high schools are classified according to their size and selectivity as well as being separated out into a separate column.

<sup>c</sup>Measure is z-scored (M=0; SD=1)
Schools’ Curriculum Structures

As discussed in Chapter 3, I conceptualize the math curriculum as a sequence of educational opportunity (Stevenson, Schiller, & Schneider, 1994) and hypothesize that how far students progress through these curricular sequences or “pipelines” will be determined by two interacting forces; first, the structure of the curriculum—both what’s offered and how students are assigned to courses; and second, students’ own decisions, motivation, ability and performance (Sorensen & Hallinan, 1977). Crucial to my theoretical framework, is the notion that the same student might experience vastly different educational experiences in different schools with different curricular structures (Sorensen, 1987; Garet & DeLany, 1988). This belief is also a critical underlying assumption of both the constrained curriculum hypothesis and small school reform more generally. If small schools concentrate their resources on offering a narrow, academic, college preparatory curriculum, all of their students will benefit, particularly disadvantaged students because this type of curricular structure eliminates the “decision points” that allow for unequal student backgrounds to manifest themselves in the form of unequal course selection (Lee, Croninger & Smith, 1997; Crosnoe & Schneider, 2010).

However, as discussed earlier, I argue that the assumption that small schools will necessarily implement this type of constrained, academic curriculum is overly simplistic. A sociological lens is needed to understand that other factors such as schools’ student body composition may play a mitigating role in how small school reform affects students’ curricular experiences in the real world. For example, because we now know that NSN schools in New York City serve a more academically underprepared student body, on average, than other NYC schools, it is possible that administrators at these schools might decide to target their curriculum to getting students through lower-level math courses at the expense of advanced course offerings. Thus, having already discussed the socio-demographic student body composition of NSN schools relative to the alternative high school
option in New York City, the next step in my investigation is to examine the curricular structures that these schools put in place.

Table 5.3 presents my findings on schools’ curricular structures and their academic composition for the 2005-06 freshman cohort (see Table A.8 in Appendix A for the 2006-07 cohort). At first glance it might look as if I have repeated the entering math and ELA achievement averages from the previous table, however there is an important distinction between how the two estimates are calculated. In the previous table I presented the average entering achievement (and the standard deviation of that achievement) for each school type, ie. the average across NSN schools versus the average across specialized schools, etc. In table 5.3, on the other hand, I present the average of the within-school means and standard deviations of entering achievement. Theoretically, the within-school measures are much more important to the curricular decisions that any given school will make. For example, a wide standard deviation of entering math achievement across all small schools could simply mean that some small schools have really low achieving students while others have high achieving students. But a wide within-school standard deviation would mean that a given school would have to deal with a freshman class with substantial variation in math skills.

**New, small schools must serve a range of math achievement levels.** As displayed in Table 5.3, I find that in 05-06 the average NSN within-school entering math achievement is a fifth of a standard deviation below the system-wide average (ES=-0.197) but that the average within-school standard deviation of entering math achievement is comparable to the average across all schools. The only school type with a noticeably higher within-school standard deviation of entering math achievement is large nonselective schools. Unsurprisingly, selective schools in each size category appear to have somewhat lower standard deviations than their nonselective counterparts. Interestingly, average within-school standard deviations in ELA achievement seem to follow a slightly different pattern, with school size appearing to be a greater indicator of within-school
variations in achievement than selectivity. Smaller schools have narrower standard deviations than larger schools and selective schools appear to have wider standard deviations than nonselective schools. However, because this is a study of the math curriculum, entering math achievement is more salient.

The system-wide within-school average entering math achievement decreases slightly between 2005-06 and 2006-07, perhaps because there are an increasing number of smaller, low achieving schools, and proportionally fewer high achieving schools. Given what we know about the decreased average entering achievement of students in NSN schools in 06-07 it is unsurprising that the average within-school entering math and ELA achievement in NSN schools declines as well, dropping from -0.197 to -0.232 standard deviations below average in math and from -.159 to -.202 standard deviations below average in ELA. This type of drop was also exhibited by VSN schools in math (but not ELA) entering achievement.

Thus, when it comes to entering math achievement, NSN schools appear to enroll students with lower average math skills than most alternative school options but not necessarily a narrower range of skills. As a result, it will be interesting to assess how students of different entering achievement levels are affected by attending a NSN school. If, as predicted, small schools cannot respond to variation in incoming academic preparedness by offering a wide range of courses, then the level at which these schools choose to target their curriculum will have important implications for their students’ curricular experiences, access to courses, and the potential interactions between student characteristics and the effect of attending a NSN school.

**New, small schools: more equitable but less excellent?** The most direct indicators of schools’ curriculum structure are how far the average student progresses through the math curriculum, and how much within-school variation there is in this progress. I find that in 05-06, the average within-school mean distance through the math curricular pipeline is 3.60, which is
equivalent to having completed the first semester of Algebra II/Trigonometry, but not finishing the sequence or beginning pre-calculus. While students in NSN schools advance slightly less far than the system-wide average (3.48), they do outperform students from large nonselective schools by over half a semester. By contrast, the average student attending a specialized high school progresses through at least the first semester of Calculus.

Interestingly, NSN schools also have below-average variation in progress through the math pipeline within schools, roughly one year compared to one and a third for large selective schools and one and a half for large nonselective schools. In other words, it appears that NSN schools might be more internally equitable when it comes to math course-taking, even if the average NSN student does not advance as far as the average student in most other types of schools. Clearly within school equity is desirable, however, just as clearly we want that equity to be at high level of competency. Schools that are equally ineffective for all students are not helping anyone. Moreover, when combined with my findings on between school stratification by race, class, and academic achievement, these findings on the internal equity of the NSN schools become less hopeful. Before jumping to conclusions, however, we must remember that these are only simple descriptive statistics. Recall from the previous tables that substantial selection of students into schools occurs based on a host of characteristics that might affect their success in the math curriculum. Conclusions on the success of NSN schools, therefore, must wait for analyses that can adjust for this selection bias.

Also presented in Table 5.3 are the within-school means and standard deviations of the distance through the math pipeline of the first math course students take in high school. As displayed in table 5.3, I find that the average first math course of students in NSN schools is again just slightly lower than the system-wide average. It is roughly equivalent, however, to the within-school average for large nonselective schools, which is interesting given that NSN schools had a
higher average distance through the math pipeline. NSN schools also have a lower average withinschool standard deviation of first math course than any other school type (except VSN schools). This figure supports the constrained curriculum hypothesis, suggesting that there is less initial curricular differentiation in small schools. Large schools, both nonselective and selective, have the widest variation in first math course taken in high school.

Between the 05-06 and 06-07 school years, the average within-school mean distance through the math pipeline of students in NSN schools decreased from 3.48 to 3.23. The substantive meaning of this drop is that by 06-07 the average NSN student is over half a semester shy of completing the first semester of Algebra II/Trigonometry. This decrease in progress through the pipeline was not caused by students beginning their first math course in high school at a lower level, because the average distance of the first math course in NSN schools did not decline between cohorts. The average within-school standard deviations of both the first and four-year distances through the math pipeline did increase slightly in NSN schools in 06-07, perhaps due to a slight widening of the within-school variation in entering math achievement.

**Students in new small schools cut-off from advanced math opportunities.** My results suggest that the largest curricular differences among school types are not in average course-taking measures such as the mean or standard deviation of students’ progress through the math pipeline, but rather in whether schools offer advanced curricular options such as Calculus or any Advanced Placement (AP) or International Baccalaureate (IB) math courses. When it comes to having the capacity to offer these types of courses, school size is clearly a limiting factor. NSN schools are the least likely of any school type to offer Calculus or AP/IB math (around a fifth of schools offer each option). Even selective small schools do not offer these courses at high rates (43.7 percent offer Calculus and 28.7 offer AP/IB math). This is in comparison to over 90 percent of large nonselective schools and 100 percent of specialized schools. Interestingly, large selective schools have somewhat
lower rates of offering these courses (80 percent offer Calculus and 68 percent offer AP/IB math) than large nonselective schools. Midsize selective schools predictably fall between small and large schools. Another substantial difference between the 05-06 and 06-07 cohorts is the percent of NSN schools offering an AP/IB math course, which decreased by half.

Clearly students cannot take courses that are not offered by their schools. Thus a high achieving student attending a NSN school might have their progress through the math curriculum artificially circumscribed by the lack of advanced course offerings. Although the disparities in advanced course offerings across schools are severe, they are not inconsistent with the extant literature. Recall from Chapter 3, a recent study of course offerings in Florida schools found that only 3% of schools in the smallest size decile (enrollments of 22 to 365) offer AP/IB math compared to 100% of schools in the top three deciles (enrollments of 2,224 and above). The study concluded that schools this small did not have a critical mass of “far above average” achieving students to warrant advanced math course offerings.

When it comes to lower level curricular options in New York City, there are less dramatic differences across school types. For example, the percentage of total courses taken below Algebra in difficulty ranges from a low of 6.21 percent in VSN schools to a high of 8.47 percent in mid-size selective high schools. NSN high schools do devote more of their courses to prep or support classes than any other school type (6.16 percent). But the differences are not substantial, with large nonselective schools offering roughly five percent prep courses. Specialized high schools offer the least number of prep courses (less than one percent).

**Summary and Implications.** The implications of these findings are deeply problematic. The average student in New York City public schools does not complete the second semester of Algebra II/Trigonometry and the average student in a NSN school does not even complete the first semester. Meanwhile, the highest achieving and most advantaged students are stratified into elite,
specialized high schools in which the average student makes it at least part of the way through Calculus. Extant literature suggests that progressing past Algebra II is a critical determinant of post-secondary participation and success (Adelman, 2006). Moreover, some research suggests that while over half of students who progress to Calculus meet college and career readiness benchmark in Math, only seven percent of students who persist only to Geometry do so (ACT, 2011). NSN schools do appear to be more internally equitable when it comes to student course-taking, yet they are equitable at a low-level of content mastery and consequently appear unequal when compared to other, more selective schools.

Even more disturbing are the disparities in the availability of advanced math course offerings across schools in New York City. Only 18% of NSN schools offer AP/IB math and only 21% offer Calculus. The patterns of advanced course offerings suggest that the availability of these courses is related to both schools’ size and selectivity. And since an increasing proportion of students in New York City are attending small, nonselective schools, it follows that an increasing proportion of students are cut off from the opportunity of choosing to take advanced math. As discussed in Chapter 3, I conceptualize student course-taking as being influenced both by schools’ curricular structures and by students’ ability and motivation. In schools where advanced math courses are not offered, student ability and motivation no longer matters—they cannot choose to take courses that are not offered. Given what we know about the low entering achievement of students in NSN schools, it may be true that the majority of these students would not make it to Calculus, let alone AP Calculus, even if these courses were offered. However, the lack of availability of these courses only underscores the need for evaluations of the impact of NSN schools to examine whether they are equally beneficial to low, middle, and high achieving students.
<table>
<thead>
<tr>
<th>Academic Composition and Curricular Structure of DOE High Schools Serving the 2005-2006 Freshman Cohort by School Size and Selectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>New, Small, Nonselective (&lt;N=103)</td>
</tr>
<tr>
<td>Entering math achievement&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
</tr>
<tr>
<td>Entering reading achievement&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
</tr>
<tr>
<td>Distance through the Math&lt;sup&gt;e&lt;/sup&gt; Curricular Pipeline (4 yr)</td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
</tr>
<tr>
<td>Distance through the Math Curricular Pipeline (1st course)&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
</tr>
<tr>
<td>School offers Calculus (%)&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>School offers AP/IB math (%)&lt;sup&gt;h&lt;/sup&gt;</td>
</tr>
<tr>
<td>Percent of credited courses below Algebra in difficulty</td>
</tr>
<tr>
<td>Percent of courses in “Prep/Support” category</td>
</tr>
</tbody>
</table>

<sup>a</sup> Curriculum structure group means are calculated on a sample that does not include students who transfer high schools so as not to attribute course-taking to the wrong school. Entering achievement group means are calculated on the full entering 9th grade class.

<sup>b</sup> See Measures section for descriptions of school size and selectivity classifications.

<sup>c</sup> Column school counts do not sum to 322 because Specialized high schools are classified according to their size and selectivity as well as being separated out into a separate column.

<sup>d</sup> Within school means and standard deviations of Eighth-grade test scores (M=0; SD=1), averaged over school size categories.

<sup>e</sup> The most advanced math course passed in four years of high school (0=No math, through 8=Advanced topics beyond Calculus)

<sup>f</sup> Level of first math course taken in high school (0=No math, through 8=Advanced topics beyond Calculus)

<sup>g</sup> School is designated as offering Calculus if three or more students from the 2005-06 cohort take an AP or IB math course.

<sup>h</sup> School is designated as offering AP or IB math if three or more students from the 2005-06 cohort complete an AP or IB math course.
Who Progresses Through the Math Pipeline?

The most advanced math course that students complete in high school is not only an indicator of their mathematics skills and academic success, it is also a signal to universities and potential employers of students’ ability and motivation and consequently serves as a filter for their subsequent college and career opportunities. Thus the next step in my descriptive analyses is to explore observed differences among students, differentiated by groups that characterize their progress through the math pipeline. I compare students’ socio-demographic and academic achievement characteristics for three groups: (1) “low”: those whose most advanced math course passed was below Algebra II/Trigonometry in difficulty [32.9 percent], (2) “middle”: those who completed both semesters of Algebra II/Trigonometry but did not reach Pre-Calculus [44.5 percent], and (3) “high”: those who reached Pre-Calculus or beyond [22.6 percent]. These descriptive statistics (displayed in Table 5.4) will reveal which students remain mired in low levels of the math curriculum and which students progress further through the pipeline to advanced, college-preparatory subjects like Calculus and Pre-Calculus (see Appendix Table A.9 for the 2006-07 cohort).

Substantial socio-demographic disparities in math course-taking. My results suggest that there are significant gender differences in progress through the mathematics pipeline. Female students are more likely than male students to reach either middle (53.1 percent) or high math levels (54.9 percent), and are less likely to fail to complete Algebra II (44.5 percent). This is consistent with extant research, which indicates that girls no longer lag behind boys in high school math coursework (Freeman, 2004). I also find substantial racial disparities in progress through the math curriculum. Black and Hispanic students are both overrepresented in the low math group (34.3 percent and 47.4 percent, respectively) and underrepresented in the high math group (20.2 percent and 22.9 percent). By contrast, white students are more likely to reach high math levels (21.9 percent) and less likely to
remain in low math levels (10.1 percent). Asian students exhibit the most dramatically positive math course-taking outcomes; they make up a full third (34.3 percent) of the high math category and only 7.29 percent of the low math category.

Poverty is also related to students’ progress through the math pipeline, particularly advancing to Pre-Calculus or above. Students who qualify for FRPL make up 66.9 percent of the high course-taking group as compared to 77.6 percent of the middle group and 79.8 percent of the low group. Students who are overage for eighth grade, who qualify for Special Education, and who are English Language Learners are all also overrepresented in the low math level group and underrepresented in the high math level group. Overage students make up 26.4 percent of the low-level group but only 14 percent of the middle group and 7.5 percent of the high group. Similarly, Special Education students make up a fifth of the low group but only 7.1 percent of the middle group and 1.3 percent of the high group. ELL students make up over 15 percent of the low group but only 8.7 percent of the high group. Students who do not make it through Algebra II also have substantially more absences and tardies in eighth grade than students who make it through the middle or high math course levels. For example low course-taking students were absent, on average, 13 more days in eighth grade than students in the high course-taking group.

Finally, one of the most important indicators of students’ success in the math curriculum is their level of mathematics achievement upon entering high school. Students in the low course-taking group have much lower entering mathematics achievement than their peers in the middle course-taking group (ES=-.587). The same is true on the opposite end of the spectrum. Students in the high course-taking group have nearly a full standard deviation higher entering math achievement than students in the middle group (ES=.945). Entering ELA achievement follows the same pattern, although the group mean differences are not quite as dramatic as for math.
Table 5.4: Group Means of Student Characteristics, Differentiated by the Highest Math Course Completed*  
(2005-06 cohort; N=48,573 students in 322 schools)

<table>
<thead>
<tr>
<th></th>
<th>Below Algebra II (N=15,980) Mean</th>
<th>Algebra II/Trig but below Pre-Calculus (N=21,639) Mean</th>
<th>Pre-Calculus or above (N=10,954) Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Female</td>
<td>45.7***</td>
<td>53.3</td>
<td>54.9*</td>
</tr>
<tr>
<td>Percent Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>7.29***</td>
<td>13.2</td>
<td>34.3***</td>
</tr>
<tr>
<td>Black</td>
<td>34.3**</td>
<td>32.5</td>
<td>20.2****</td>
</tr>
<tr>
<td>Hispanic</td>
<td>47.4***</td>
<td>39.5</td>
<td>22.9****</td>
</tr>
<tr>
<td>Native/Multi</td>
<td>0.7</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>White</td>
<td>10.1***</td>
<td>13.7</td>
<td>21.9***</td>
</tr>
<tr>
<td>Percent eligible for free and reduced-price lunch</td>
<td>79.8***</td>
<td>77.6</td>
<td>66.9***</td>
</tr>
<tr>
<td>Percent overage for eighth grade</td>
<td>26.4***</td>
<td>14.0</td>
<td>7.5***</td>
</tr>
<tr>
<td>Percent Special education</td>
<td>19.6***</td>
<td>7.1</td>
<td>1.3***</td>
</tr>
<tr>
<td>Percent English Language Learner</td>
<td>15.3***</td>
<td>13.3</td>
<td>8.7***</td>
</tr>
<tr>
<td>Days absent in eighth grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>20.9***</td>
<td>12.0</td>
<td>7.8****</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(18.9)</td>
<td>(9.45)</td>
<td>(6.97)</td>
</tr>
<tr>
<td>Days tardy in eighth grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.2***</td>
<td>12.5</td>
<td>6.7****</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(24.1)</td>
<td>(17.9)</td>
<td>(12.7)</td>
</tr>
<tr>
<td>Entering math achievementa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-.621***</td>
<td>-.034</td>
<td>.911***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(.938)</td>
<td>(.766)</td>
<td>(.778)</td>
</tr>
<tr>
<td>Entering ELA achievementa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-.530***</td>
<td>-.055</td>
<td>.800***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(.860)</td>
<td>(.801)</td>
<td>(.998)</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01, ***p<.001; For significance testing, Below Algebra II (column 1) and Pre-Calculus and above (column 3) are compared to Algebra II but below Pre-Calculus (column 2).  
*aMeasure is z-scored (M=0; SD=1)
School Characteristics Associated with Greater Student Progress through the Math Curriculum

What distinguishes schools with different levels of average student progress through the math curriculum? In addition to substantial variation in mathematics course-taking between students, there is also variation in average student course-taking between schools. Similar to exploring observed differences in groups of students differentiated by their most advanced math course, it is also helpful to examine the differences in school characteristics exhibited by groups of schools differentiated by the average progress of their students through the math curriculum. Thus, I compare school demographics and curriculum structure for three group of schools: (1) those with low average math course-taking (more than one standard deviation below the mean), (2) those with middle average course-taking (within one standard deviation of the mean), and (3) those with high average course-taking (more than one standard deviation above the mean). Table 5.5 presents the results for the 2005-06 cohort (see Appendix Table A.10 for the 2006-07 cohort).

For these analyses I separate school size, selectivity, and newness into separate measures to get a better sense of which characteristics are most associated with average progress through the math pipeline. Note that given the much smaller sample size (322 schools) some group mean differences will be non-significant simply due to a lack of statistical power. Thus although none of the school size categories are significantly related to average course-taking, it is interesting to note that small and mid-size schools make up higher proportions of the high course-taking group than the middle or low groups, while the reverse is true for large schools. Again, specialized high schools demonstrate their elite level of academic success as they all rank in the high average course-taking group. But even beyond specialized high schools, academic selectivity is associated with average progress in mathematics. Academically non-selective schools are underrepresented in the high average course-taking group (27.7 percent compared to 52.6 percent in the middle group). Whether a school is “new” (2003-03 or later) does not appear to be related to average student course-taking.
NSN schools make up a slightly higher percentage of the low course-taking group than the middle or high groups, but the differences are not statistically significant.

When it comes to student body composition, the aggregate patterns are consistent with the individual predictors of progress. Schools in the low average course-taking group have higher proportions of black and Hispanic students (92.7 percent) compared to middle average course-taking schools (85.2 percent) and particularly compared to high average course-taking schools (56.4 percent). Similarly, high average course-taking schools have lower percentages of free or reduced price lunch eligible students, overage students, ELL students, and special education students than schools with middle or low average course-taking.

Unsurprisingly, school-average entering student achievement in both math and ELA is strongly related to school-average student progress through the math curriculum. The most dramatic differences are between the low and high average course-taking groups. High average course-taking schools enroll students with average entering math and ELA achievement that is a full standard deviation higher than low average course-taking schools (ES=1.06 for math and ES=1.07 for ELA). High average course-taking schools also have a wider standard deviation in entering math achievement than either middle or low average course-taking schools. Moreover, between the 05-06 and 06-07 cohorts, the disparities in achievement between low and high average course-taking schools rose to 1.28 standard deviations for math and 1.20 standard deviations for reading.

Now turning to the curriculum structure measures, students in high average course-taking schools also begin their high school math careers in more advanced math courses than their peers in middle and low average course-taking schools. However it is interesting to note that while there are significant differences in average difficulty of students’ first math course, these differences are not as large as for the most advanced math course. In other words, while students in high average course-taking schools begin farther ahead, they also make larger gains, on average, during high school than
their peers in middle or low average course-taking schools. Mirroring their wider within-school variation in entering math achievement, high average course-taking schools also have a wider standard deviation of the first math course taken than either middle or low average course-taking schools.

When it comes to offering advanced math courses, there are also substantial differences across groups. Over 90 percent of high average course-taking high schools offer Calculus compared to 46.2 percent of middle average course-taking schools and only 26.8 percent of low average course-taking schools. A similar pattern holds for AP/IB courses; over three quarters of high average course-taking schools, but only roughly a third of middle average and a quarter of low average course-taking schools offer any AP or IB math courses. On the other hand, when it comes to lower level math courses, low average course-taking schools devote roughly twice as much of their curriculum to courses below Algebra in difficulty and to prep or support courses than high average course-taking schools.
Table 5.5: Structural and Compositional Characteristics of High Schools, Differentiated by the School Average of Students’ Highest Math Course Completed\( ^a \) (2005-06 cohort; \( N=322 \))

<table>
<thead>
<tr>
<th>School Structure</th>
<th>Low average course-taking( ^b ) (N=41)</th>
<th>Middle average course-taking( ^c ) (N=234)</th>
<th>High average course-taking( ^d ) (N=47)</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Size(%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small( ^e )</td>
<td>56.1</td>
<td>64.5</td>
<td>63.8</td>
</tr>
<tr>
<td>Mid-Size</td>
<td>14.6</td>
<td>14.5</td>
<td>19.2</td>
</tr>
<tr>
<td>Large</td>
<td>29.3</td>
<td>20.9</td>
<td>17.0</td>
</tr>
<tr>
<td>Specialized</td>
<td>0.0</td>
<td>0.0</td>
<td>17.0***</td>
</tr>
<tr>
<td>Academically non-selective (%)</td>
<td>70.7</td>
<td>52.6</td>
<td>27.7**</td>
</tr>
<tr>
<td>New School (%)</td>
<td>41.5</td>
<td>47.0</td>
<td>36.2</td>
</tr>
<tr>
<td>New<em>Small</em>Nonselective (%)</td>
<td>39.0</td>
<td>32.9</td>
<td>21.3</td>
</tr>
</tbody>
</table>

**School Composition**

<table>
<thead>
<tr>
<th>Percent Black and Hispanic</th>
<th>92.7</th>
<th>85.2</th>
<th>56.4***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Free or Reduced Price Lunch Eligible</td>
<td>69.3</td>
<td>72.8</td>
<td>51.0***</td>
</tr>
<tr>
<td>Percent average for 8(^{th} ) grade</td>
<td>26.4***</td>
<td>19.6</td>
<td>9.42***</td>
</tr>
<tr>
<td>Percent English Language Learner</td>
<td>11.0</td>
<td>13.9</td>
<td>6.97*</td>
</tr>
<tr>
<td>Percent Special Education</td>
<td>17.6***</td>
<td>10.7</td>
<td>5.22***</td>
</tr>
</tbody>
</table>

Entering math achievement\( ^f \)

<table>
<thead>
<tr>
<th>Average w/in school mean</th>
<th>-.470****</th>
<th>-.177</th>
<th>.589***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average w/in school std. dev.</td>
<td>.861</td>
<td>.807</td>
<td>.657***</td>
</tr>
</tbody>
</table>

Entering reading achievement\( ^f \)

<table>
<thead>
<tr>
<th>Average w/in school mean</th>
<th>-.436**</th>
<th>-.202</th>
<th>.633***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average w/in school std. dev.</td>
<td>.763</td>
<td>.768</td>
<td>.823</td>
</tr>
</tbody>
</table>

**Curriculum Structure**

Distance through the Math Curricular Pipeline (4 yr)\( ^g \)

<table>
<thead>
<tr>
<th>Average w/in school mean</th>
<th>2.61***</th>
<th>3.49</th>
<th>5.00***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average w/in school std. dev.</td>
<td>1.14</td>
<td>1.16</td>
<td>1.32*</td>
</tr>
</tbody>
</table>

First Math Course\( ^h \)

<table>
<thead>
<tr>
<th>Average w/in school mean</th>
<th>1.20*</th>
<th>1.36</th>
<th>1.71***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average w/in school std. dev.</td>
<td>.296</td>
<td>.311</td>
<td>.429**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School offers Calculus (%)</th>
<th>26.8*</th>
<th>46.2</th>
<th>93.6***</th>
</tr>
</thead>
<tbody>
<tr>
<td>School offers AP/IB math (%)</td>
<td>26.9</td>
<td>36.8</td>
<td>78.7***</td>
</tr>
<tr>
<td>Percent of credited courses below Algebra in difficulty</td>
<td>11.6*</td>
<td>7.30</td>
<td>5.89</td>
</tr>
<tr>
<td>Percent of courses in “Prep/Support” category</td>
<td>6.93</td>
<td>5.31</td>
<td>2.48**</td>
</tr>
</tbody>
</table>

\*\( p<.05, \)**\( p<.01, \)**\( p<.001; \) For significance testing, low average course-taking and high average course-taking groups are compared to middle average course-taking.

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

\( ^b \) In schools with low average course-taking, the most advanced math course completed by the average student was below Algebra II in difficulty.

\( ^c \) In schools with medium average course-taking, the most advanced math course completed by the average student was Algebra II (first or second term).

\( ^d \) In schools with high average course-taking, the most advanced math course completed by the average student was beyond Algebra II in difficulty.

\( ^e \) See Measures section for descriptions of school size and selectivity classifications.

\( ^f \) Within school means and standard deviations of Eighth-grade test scores (\( M=0, SD=1 \)), averaged over school size categories.

\( ^g \) The most advanced math course passed in four years (0=No math, through 8=Advanced topics beyond Calculus).

\( ^h \) Level of first math course taken in high school (0=No math, through 9=Advanced topics beyond Calculus).
Summary

The results of my descriptive analyses confirm that NSN schools do enroll a more disadvantaged student body, on average, than the alternative public high school options in New York City. These findings also situate NSN schools in a complex high school landscape characterized by a great deal of sorting and stratification of students between schools. Specialized high schools represent the epitome of this type of student selection into schools and often the disparities in student body composition, curriculum structure, and academic outcomes between these elite schools and the NSN schools were quite severe. When it comes to curriculum structures, I find that NSN schools have slightly lower average progress through the math pipeline, but that they also have a slightly narrower standard deviation of progress, despite enrolling freshmen whose levels of math skills vary to roughly the same degree as most other high school options. NSN schools are also much less likely than larger schools in NYC to offer Calculus or AP/IB math courses. Indeed, it appears that the majority of NSN schools completely cut their students off from the opportunity of taking these advanced math courses. Thus at first glance, it appears that NSN schools may be more “equitable”—at least internally—but not more “excellent” than other high schools options in New York City. However, given the documented degree of student selection into schools, I argue that analytic models that address selection bias are needed before a conclusion on the impact of NSN schools can be drawn.

In this chapter I also detail the characteristics of both students and schools that are associated with high progress through the math curriculum. Unsurprisingly, many student socio-demographic and academic characteristics were strongly related to progress through the math curriculum. These significant relationships only underscore the need to compare apples to apples when assessing the impact of NSN schools. When it comes to school characteristics, first impressions indicate that student body composition, the academic selectivity schools, and their
curricular structures might be more strongly associated with average progress through the math curriculum than school size. In the next chapter I will use multivariate, multilevel models to better describe the associations between school factors such as size and selectivity and students’ progress through the math curriculum, controlling for students’ socio-demographic and academic characteristics. I will also test whether attending a NSN school may be more beneficial for certain types of students, as the extant literature suggests. Next, I will use multiple propensity score matching techniques to account for the clear threat of selection bias and estimate the effect of attending a NSN school on students’ progress through the math curriculum. And finally, I will utilize students’ course grades and Regents exam scores to assess the rigor of courses covering the same subject matter across schools. This will provide insight into whether the math pipeline measure is a reasonable indicator of students’ content mastery in NSN schools, relative to the other high school options in New York City.
Chapter 6

“Constrained Downwards”: New Small Schools Foster Equity but not Excellence

The descriptive results presented in the previous chapter begin to paint a picture of the newly created small high schools and where they fit in the broader landscape of public secondary school options in New York City. These schools are geographically located in some of the most disadvantaged parts of the five boroughs and serve a student population that is lower income, lower achieving, and more heavily black and Hispanic than any of the other high school alternatives. With regards to average progression through the math curriculum, the NSN schools rank near the bottom. However, they do appear to out-perform large nonselective schools, at least in the 2005-06 cohort. In this chapter, I will go beyond simple descriptive statistics to assess how well NSN schools are performing after adjusting for the fact that they enroll a more underserved student body than almost all other high school options in the city.

Perhaps the most startling finding from the descriptive results is the intensity of academic stratification—both in terms of inputs and outcomes—across high schools in New York City. In specialized high schools, an elite group of the highest achieving eighth graders, who are disproportionately white, Asian, and higher income, follow an advanced mathematics program in which the average student reaches Calculus. Meanwhile, nonselective schools enroll initially low-achieving, low-income students, who are disproportionately black and Hispanic, English language learners, overage for grade, and who have higher rates of absences and tardies. Many of these schools do not offer Calculus or any type of AP/IB math. The average student in these schools will not even make it through Algebra II.

The high degree of academic stratification in New York City has several important implications for my assessment of small high school reform. First, from a methodological perspective, the intensity of the sorting of students across high schools underscores the need to use
sophisticated statistical techniques to ensure that I am truly comparing “apples to apples” when estimating the small school effect on student course-taking. Failing to do so would unfairly penalize NSN schools for serving a more disadvantaged, initially lower achieving student body and unfairly reward academically selective high schools for enrolling more privileged, higher achieving students.

Second, from a theoretical perspective, a high degree of stratification may limit the ability of small schools to really level the playing field. At a systemic level, the equalizing capacity of small school reform lies primarily in the elimination of tracking and its replacement with a core academic curriculum that all students take. While such a reform may address issues related to the stratification of students between tracks within schools, it does not solve—and may even exacerbate—the issue of stratification between schools themselves. To the extent that stratification occurs between schools in a given district, increasing the within-school equity in course-taking by reducing school size and constraining the curriculum may have a limited impact on the system-wide level of equality.

In this type of system, where NSN schools serve predominantly disadvantaged, academically underprepared students while the most privileged, academically advanced students sequester themselves in a handful of selective high schools, small school reform can no longer leverage the equalizing potential of integrating students day-to-day educational experiences. Rather, the success of the reform hinges completely on whether small schools can elevate disadvantaged students’ educational experiences and outcomes within their location at the bottom of a stratified system.

We know from the descriptive results presented in Chapter Five that small school reform did not succeed at the herculean task of equalizing educational outcomes in New York City, yet it is still possible that these new, small schools may be an improvement over the alternative high school options that serve similarly disadvantaged students. According to the extant literature, small schools are hypothesized to be particularly advantageous for low-income, minority, and low-achieving students. Thus, while small school reform in New York City may not have leveled the playing field,
the new, small, nonselective schools might still be outperforming the type of large, nonselective high schools that they replaced. This in itself would be a substantial accomplishment.

In this chapter, I utilize both multivariate, multilevel regression and propensity score matching techniques to assess whether students in the NSN schools in New York City progress farther through the math curriculum than their peers in the alternative public high school options, after controlling for their socio-demographic and academic characteristics. I also investigate whether the NSN schools are equally effective for all types of students. Is it true that these schools are more beneficial for the most disadvantaged students? And what of critics’ fears that the new small schools would do a disservice to high achieving students by cutting them off from advanced coursework? I will also explore whether the NSN effect is the same for both the 2005-06 and 2006-07 cohorts. Finally, I will utilize students’ course grades and Regents exams scores to explore whether the rigor of math courses varies across schools in New York City. Specifically, I will test whether the rigor of courses in NSN schools is comparable to the other high school options.

**Multivariate and Multilevel Descriptive Results**

**Unconditional model**

The first step in multilevel modeling is to partition the variance in the outcome into its within- and between-school components. It is only the portion of the total variance that lies between schools that can be modeled as a function of school factors such as size. Table 6.1 displays the psychometric properties of my outcome: the highest mathematics course completed. These results were computed as a fully unconditional multilevel model, meaning that no predictors were specified at either the student- or school-level. The intraclass correlation (ICC) represents the proportion of the variability in students’ math course-taking that is systematically between schools. The estimated ICC of .223 indicates that approximately 22 percent of the variability in progression through the math curriculum lies between schools. In other words, despite the between-school academic
stratification that occurs in New York City, over three-quarters of the variance in progression through the math pipeline occurs *within* schools. Within the school effects literature, an ICC of this level is relatively common with academic outcomes.

| Table 6.1: Multilevel Psychometric Properties of the Math Pipeline Measure (N=101,261 students in 342 schools) |
|-------------------------------------------------|----------------------|
| Math Pipeline (highest mathematics course completed) |  |
| Pooled within-school variance ($\sigma^2$) | 1.96  |
| Between-school variance ($\tau$) | .565  |
| Intraclass Correlation (ICC) | .223  |

$^a$ ICC = $\tau / (\tau + \sigma^2)$

**Student-level models**

Table 6.2 displays the results of my analyses of the student characteristics associated with progression through the mathematics pipeline. Unlike the results from the previous chapter, in these multivariate models I am able to adjust the associations of students’ socio-demographic and academic characteristics for one another. This allows me to assess, for example, whether racial gaps in mathematics course-taking remain even after controlling for attributes such as students’ initial math achievement, poverty, and ELL status. To maintain consistency and ease of interpretation, the centering of independent variables is consistent across all multilevel models. Race/ethnicity indicators and students’ entering mathematics achievement are the only group-mean centered measures, meaning that these estimates specifically compare students within the same schools.

In Model 1, I establish the relationships between students’ socio-demographic characteristics and the most advanced math course passed. These associations are adjusted for one another but not yet adjusted for students’ behavioral or academic characteristics. Even after controlling for gender and FRPL, ELL, and Special Education statuses, substantial racial/ethnic disparities in progression through the math pipeline remain. During their tenure in high school, Asian students advance roughly two-fifths of a year farther in math than their white counterparts attending the same
schools. By contrast, black and Hispanic students both lag behind their white peers by over a third of a year in math course-taking. Of the five racial/ethnic groups, Hispanic students progress the least far through the math curriculum in NYC. Students in the Native American/Multi-ethnic category progress roughly one-sixth of a year less far through the math pipeline than white students.

Female students advance nearly one-fifth of a year farther than males in math, suggesting that the gender gap of old has reversed when it comes to math course-taking in high school. Unsurprisingly, Special Education and ELL status are both negatively associated with progress in math. FRPL eligibility is surprisingly non-significant in Model 1 and even becomes slightly positive and significant in later models, reinforcing the concern that this is a very weak measure of students’ socioeconomic status. Finally, it is interesting to note that overall, students in the 2006-07 cohort advance slightly less far in math than their counterparts in the 2005-06 cohort.

In Model 2, I further adjust for students’ behavioral and academic characteristics prior to entering high school. Predictably, entering math achievement is very strongly related to eventual progress through the math pipeline. Within schools, a one standard deviation increase in entering math achievement is associated with progressing nearly a semester farther through the math curriculum. After adjusting for math achievement as well as the other academic and behavioral attributes in Model 2, the racial disparities in math course-taking are reduced but do not disappear. Controlling for all other covariates, Asian students still advance nearly half a semester farther in math than white students, who in turn progress one-fifth of a year farther than Hispanic students in the same school. Meanwhile, the within-school black-white gap in math course-taking is reduced to one-tenth of a year.

In Model 3, I additionally control for the first math course taken in high school. Which math course students begin high school with is an indicator of both their prior experience and success in the math curriculum in middle school as well as the curriculum structure and assignment
mechanisms in place in their high school. 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, even after controlling for entering math achievement and all other student background characteristics. Racial disparities in advancement through the math curriculum are slightly reduced after controlling for first math course but remain significant, indicating that they are not purely a function of where students are when they enter high school but what goes on during high school as well. Put another way, even after adjusting for where they start, Asian students progress farther in the math curriculum than white students, who in turn progress farther than black and Hispanic students in the same schools.

Finally, in Model 4 I interact first math course taken in high school with all other student-level covariates\textsuperscript{12}. A negative interaction with first math course indicates that the link between which math course students start in and which math course they finish in is weaker for that group of students than for the comparison group. For example, the negative interaction between first math course and female students suggests that the positive association between first math course and eventual progress through the math curriculum is slightly weaker for females than it is for males. The same is true for ELLs, students in the 2006-07 cohort, and students who are overage, more mobile, and have more absences in middle school. By contrast, the interactions between entering math and ELA achievement and first math course are positive, suggesting that the relationship between first math course attempted and eventual progress through the math pipeline is even stronger for students with higher initial achievement. In other words, students with stronger entering skills are better able to capitalize on a good first math course placement than students with weaker entering skills.

\textsuperscript{12} Non-significant interactions were removed in the interest of model parsimony.
Table 6.2: Student-level Multilevel Models of Highest Math Course Completed* (2005-06 and 2006-07 Cohorts; N=101,261 students in 342 schools)

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Socio-Demographic Characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity(^b,c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>.409***</td>
<td>.226***</td>
<td>.222***</td>
<td>.221***</td>
</tr>
<tr>
<td>Black</td>
<td>-.344***</td>
<td>-.107***</td>
<td>-.089***</td>
<td>-.085***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-.369***</td>
<td>-.211***</td>
<td>-.193***</td>
<td>-.191***</td>
</tr>
<tr>
<td>Native/Multi-Ethnic</td>
<td>-.179**</td>
<td>-.079</td>
<td>-.066</td>
<td>-.058</td>
</tr>
<tr>
<td>Female</td>
<td>.194***</td>
<td>.158***</td>
<td>.157***</td>
<td>.156***</td>
</tr>
<tr>
<td>Free or Reduced Price Lunch Eligible</td>
<td>.012</td>
<td>.049***</td>
<td>.055***</td>
<td>.055***</td>
</tr>
<tr>
<td>Special Education</td>
<td>-.883***</td>
<td>.002</td>
<td>-.013</td>
<td>-.036*</td>
</tr>
<tr>
<td>English Language Learner</td>
<td>-.172***</td>
<td>.054***</td>
<td>.055***</td>
<td>.042**</td>
</tr>
<tr>
<td>2006-07 Cohort</td>
<td>-.090***</td>
<td>-.042*</td>
<td>-.066***</td>
<td>-.073***</td>
</tr>
<tr>
<td><strong>Behavioral/Academic Characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overage for 8th grade</td>
<td>-.267***</td>
<td>-.290***</td>
<td>-.294***</td>
<td></td>
</tr>
<tr>
<td>Middle school mobility</td>
<td>-.066***</td>
<td>-.067***</td>
<td>-.068***</td>
<td></td>
</tr>
<tr>
<td>Entering math achievement(^c,d)</td>
<td>.480***</td>
<td>.439***</td>
<td>.450***</td>
<td></td>
</tr>
<tr>
<td>Entering ELA achievement(^d)</td>
<td>.127***</td>
<td>.095***</td>
<td>.091***</td>
<td></td>
</tr>
<tr>
<td>Days absent in 8th grade</td>
<td>-.025***</td>
<td>-.025***</td>
<td>-.025***</td>
<td></td>
</tr>
<tr>
<td>Days tardy in 8th grade</td>
<td>-.007***</td>
<td>-.007***</td>
<td>-.007***</td>
<td></td>
</tr>
<tr>
<td><strong>First Math Course(^e)</strong></td>
<td></td>
<td>.436***</td>
<td>.390***</td>
<td></td>
</tr>
<tr>
<td><strong>Interactions:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First*Female</td>
<td></td>
<td></td>
<td>-.058***</td>
<td></td>
</tr>
<tr>
<td>First*ELL</td>
<td></td>
<td></td>
<td>-.117***</td>
<td></td>
</tr>
<tr>
<td>First*Cohort 06-07</td>
<td></td>
<td></td>
<td>-.090**</td>
<td></td>
</tr>
<tr>
<td>First*Overage</td>
<td></td>
<td></td>
<td>-.051**</td>
<td></td>
</tr>
<tr>
<td>First*Mobility</td>
<td></td>
<td></td>
<td>-.032*</td>
<td></td>
</tr>
<tr>
<td>First*Math</td>
<td></td>
<td></td>
<td>.113***</td>
<td></td>
</tr>
<tr>
<td>First*ELA</td>
<td></td>
<td></td>
<td>.047***</td>
<td></td>
</tr>
<tr>
<td>First*Days absent</td>
<td></td>
<td></td>
<td>-.004**</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.55***</td>
<td>3.60***</td>
<td>3.62***</td>
<td>3.60***</td>
</tr>
</tbody>
</table>

\(^*p<.05, **p<.01, ***p<.001\)
\(^a\) The most advanced math course passed in four years of high school (0=No math, through 8=Advanced topics beyond Calculus)
\(^b\) All racial/ethnic groups are compared to white students.
\(^c\) Measure is group-mean centered
\(^d\) 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)
**Between-school models**

*School factors associated with progress in math.* Next, I turn to the between-school multilevel models. First, I explore the relationships between school characteristics such as curriculum structure and student body composition, and students’ progress through the math curriculum (see Table 6.3). In this model, I disentangle the associations between school size, selectivity, and era of founding from one another. Then, in Table 6.4, I present the first estimate of the relationship between attending a NSN school and students’ progress through the math pipeline, controlling for students’ background characteristics. Recall from Chapter 4 however, that these estimates are still considered descriptive, and are not interpreted causally. Despite controlling for all available student background measures, there is still a possibility of omitted variable bias given the high degree of student selection into high schools in New York City. I reserve causal interpretation for my propensity score matched results (and even those are interpreted cautiously).

I present the relationships between school characteristics and students’ most advanced math course completed in Table 6.3. These associations are adjusted for one another as well as for all student background characteristics. After controlling for all covariates, I find a negative relationship between attending a large high school relative to a midsize high school, but no relationship between attending a small high school relative to a midsize high school on students’ progress through the math curriculum. Net of all other school and student characteristics, students in large high schools appear to advance roughly a third of a year less far in math than their peers in midsize schools. As a reminder, this result may flow partially from the fact that large high schools have higher rates of students dropping out than smaller schools and drop-outs are retained in my sample. My results also suggest a negative association between attending a non-selective high school and students’ advancement in math. Students in non-selective schools progress nearly one-sixth of a year less far through the math pipeline than their counterparts in selective schools.
I also find several significant associations between the attributes of schools’ curriculum structure and students’ success in math. For example, students that attend schools that offer Calculus and Advanced Placement (AP) or International Baccalaureate (IB) math courses advance farther through the math curriculum than students who do not. As discussed in more depth in my theoretical framework, schools’ course offerings can place real constraints on students’ ability to access advanced math. Students can only take what is offered, so if their school does not have the capacity to offer Calculus, then high achieving students’ true potential to progress through the math pipeline may be artificially constrained by a sort of “curricular ceiling effect”. This finding has implications for many NSN students because, as we know from Chapter 5, NSN schools are much less likely to offer these advanced course offerings than any other school type in New York City.

By contrast to the positive association with advanced courses, I find a slight negative association between the percent of courses below Algebra in difficulty in a students’ school and their progress in math. A 20 percent increase in the percent of courses below Algebra is associated with students progressing approximately one-eighth of a year less far through the math pipeline. Although the magnitude of the association is not substantial, this finding does lend some support to the constrained curriculum hypothesis, which suggests that eliminating this type of lower level course should have a positive influence on all students’ course-taking and an even stronger positive influence on disadvantaged students.

On the other hand, I also find a positive relationship between the within-school standard deviation of the most advanced math course taken and students’ progress in math. If this can be viewed as a measure of within-school curricular differentiation, then the positive coefficient indicates that, on average, students in more differentiated schools progress slightly farther through the pipeline than students in more constrained schools. This does not support the constrained curriculum hypothesis, however it is possible that I have not fully controlled for student selection
into schools. Another possibility is that schools with narrow, constrained curriculums in New York City are more likely to target their courses at a lower level of difficulty, artificially constraining the average student’s progress in math to be lower than it would be if allowed to freely manifest itself in a broader curriculum with more advanced course offerings.

Finally, when it comes to student body composition, I find a negative relationship between the percent of a school’s student body that are black or Hispanic and students’ progress in math. Even after controlling for students’ own race and all other covariates, a 20 percent increase in the percent of black and Hispanic students in a school is associated with students advancing one-fifth of a year less far in math. The reverse is true for entering student achievement. My results suggest that a one standard deviation increase in a school’s mean entering math achievement is associated with students progressing one-fifth of a year farther through the math pipeline, even after adjusting for their own achievement and all other covariates. These results indicate that school compositional factors are related to individual students’ progress in math. This lends support to my conceptual argument that other school factors such as student body composition and the level of stratification between schools may mitigate how the constrained curriculum hypothesis plays out in the context of small school reform in New York City.
Table 6.3: Between-school, Multilevel Models Exploring the Relationships Between School Factors and Students’ Progress Through the Mathematics Curricular Pipeline (2005-06 and 2006-07 cohorts; N=101,261 students in 342 schools)

<table>
<thead>
<tr>
<th>School Type:</th>
<th>Model 1b</th>
</tr>
</thead>
<tbody>
<tr>
<td>School sizec</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>.056</td>
</tr>
<tr>
<td>Large</td>
<td>-.339***</td>
</tr>
<tr>
<td>Academically non-selective</td>
<td>-.162***</td>
</tr>
<tr>
<td>New school</td>
<td>.170**</td>
</tr>
</tbody>
</table>

Curriculum Structure

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>School offers Calculus</td>
<td>.134***</td>
</tr>
<tr>
<td>School offers AP/IB math</td>
<td>.149***</td>
</tr>
<tr>
<td>Percent of credited courses below Algebra in difficulty</td>
<td>-.006***</td>
</tr>
<tr>
<td>Within-school standard deviation of most advanced math course taken (4 yr)</td>
<td>.097**</td>
</tr>
</tbody>
</table>

School Composition

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Black and Hispanic</td>
<td>-.010***</td>
</tr>
<tr>
<td>Within-school mean of entering math achievement</td>
<td>.193***</td>
</tr>
</tbody>
</table>

Constant 4.68***

*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 Model is also adjusted for all student characteristics included in Table 6.2, Model 2. The following non-significant school-level predictors were removed in the interest of model parsimony: borough indicators, the percent of students that were free or reduced price eligible, special education, overage, ELL, and the percent of total courses classified as prep/support.
c School size categories are compared to midsize schools.

New, Small, Non-selective schools. After looking broadly at the characteristics of students and schools associated with progress through the math curriculum, I now turn to the focus of my study—NSN schools. The results presented in Table 6.4 constitute the findings on the NSN effect on student course-taking that researchers would find if they only utilized multilevel regression. I argue, however, that regression controlling for only the available student-level covariates is insufficient to eliminate selection bias in the New York City high school context. Thus, after presenting the findings from my multilevel regressions, I compare them to several propensity score matching techniques.
First, in Model 1 of Table 6.4 I establish the unadjusted relationship between attending a NSN school and students’ progress through the math pipeline. Without controlling for any student or school characteristics, students in NSN schools advance over half a semester less far in math than their counterparts in the alternative public high school options in New York City. In Model 2, I adjust the NSN school effect for all student socio-demographic, academic, and behavioral characteristics as well as cohort. After controlling for these covariates, students attending a NSN school still progress roughly one-seventh of a year less far through the math curriculum than students in other public high schools. Of all the multilevel regression estimates, this is the most comparable to the results that I will present from the propensity score matching analyses because it controls for only pre-treatment characteristics of students.

Benefits depend on student background. In the next three models (3-5) of Table 6.4, I explore whether the effect of attending a NSN school may vary depending on student characteristics such as race, FRPL eligibility, and entering math achievement. I also test whether students in the 2006-07 cohort have a different experience in NSN schools than students in the 2005-06 cohort. I find that there are differential effects by cohort, for black and Hispanic students, and for students with different levels of entering math achievement. The interaction between FRPL eligibility and attending a NSN school was non-significant, which is unsurprising given the weakness of that measure in these data. First, in Model 4, I introduce the cohort by NSN school interaction. I find that, after controlling for all other student and school covariates, the negative association between attending a NSN school and progress through the math pipeline is slightly stronger for students in the 2006-07 cohort. Furthermore, after adjusting for the interaction between attending a NSN school and being a member of the 06-07 cohort, the negative coefficient for attending a NSN school becomes non-significant, meaning that for students in the 05-06 cohort there is no effect of attending a NSN school.
There are several explanations for why there might be a differential NSN effect by cohort. From the descriptive analyses in Chapter 5, we know that the average entering math achievement in NSN schools decreased more between cohorts than it did in most other school types. If some residual selection bias remains, then this demographic difference might explain the negative interaction. Alternatively, as I will discuss more fully in Chapter 7, there are policy reasons why NSN schools might perform worse in later cohorts, such as the expiration of start-up exemptions or the fading of staff energy associated with the start of a new reform. The propensity score matching analyses should provide insight into whether this interaction is due to selection bias.

Next, in Model 5 I find that the effect of attending a NSN school on student course-taking may depend on students’ race. The within-school black-white and Hispanic-white achievement gaps are completely eliminated, if not reversed in NSN schools (see Figure 6.1). This finding is consistent with the extant literature on small schools, which found more racial/ethnic equity in learning gains in smaller high schools (Lee & Smith, 1997). However, recall from the previous chapter that over 90 percent of the students in NSN schools are black or Hispanic. It is likely that the few white students attending NSN schools are different in many measured and unmeasured ways from the overall white student population in New York City. Finally, in Model 6, I examine whether the relationship between attending a NSN school and success in the math curriculum varies by students’ entering math achievement. I find that while students with very low entering achievement may benefit from attending a NSN school, as students’ entering achievement increases, the benefits they accrue from attending a NSN school diminish and eventually reverse such that middle- to high-achieving students are worse off attending a NSN school than an alternate high school option (see Figure 6.2).

These findings on the differential impact of attending a NSN school for students with different racial/ethnic and academic backgrounds is crucial to understanding both the mechanisms through which the new small schools are influencing student course-taking and the policy...
implications for the scalability of the reform. As I will discuss more in Chapter 7, these results are consistent with my conceptual framework, which emphasizes the interaction between what students bring to the table in terms of ability, preparedness, and motivation, and what curriculum structure schools put in place. The implication of this framework is that the exact same school, with the exact same curricular structure can simultaneously improve some students’ educational trajectories while harming the trajectories of others. The results presented in Table 6.4 suggest that NSN schools are doing exactly this—they are improving the educational trajectories of the lowest achievers but harming the trajectories of the middle and high achievers.

| Table 6.4: Between-school, Multilevel Models Exploring the Relationship Between Attending a New, Small, Nonselective School and Progress Through the Mathematics Curricular Pipeline* (2005-06 and 2006-07 cohorts; N=101,261 students in 342 schools) |
|---------------------------------|----------------|----------------|----------------|----------------|----------------|
|                                | Model 1       | Model 2       | Model 3       | Model 4       | Model 5       |
|                                | NSN School    | + Student    | + Cohort      | + Race        | + Achievement  |
|                                |               | Characteristics | by NSN        | by NSN        | by NSN        |
|                                |               |                | Interaction   | Interactions  | Interactions  |
| New, Small, Nonselective (NSN) School | -.270**       | -.146*         | -.118         | -.146*         | -.145*         |
| Student*NSN Interactions:      |               |                |               |               |               |
| 2006-07 Cohort                 |               | -.042*         | -.009         | -.042***       | -.042*         |
| Black*c, d                     |               | -.107***       | -.107***      | -.120***       | -.107***       |
| Hispanic*c, d                  |               | -.212***       | -.211***      | -.228***       | -.212***       |
| Entering math achievementd, e  |               | .480***        | .480***       | .481***        | .534***        |
| NSN*Cohort 06-07               |               |                | -1.18**       |               |               |
| NSN*Black*f                    |               |                |               | .209***        |               |
| NSN*Hispanic*f                 |               |                |               | .219***        |               |
| NSN*Math                       |               |                |               |               | -.171***       |
| Constant                       | 3.66***       | 3.68***        | 3.68***       | 3.68***        | 3.68***        |

*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 Model is also adjusted for all student characteristics included in Table 6.2, Model 2.
c All racial/ethnic groups are compared to whites. Asian and Native American/Multi-ethnic indicator variables and interactions are included in the models to preserve white students as the comparison group but are omitted from the table because they are non-significant.
d Measure is group-mean centered
e Measure is z-scored (M=0; SD=1)
f Interactions with non-significant racial/ethnic groups are omitted from the table but remain in the model to preserve white students as the comparison group.
Figure 6.1: The relationship between race/ethnicity, attending a NSN school, and math course-taking

Figure 6.2: The relationship between entering math achievement, attending a NSN school, and math course-taking
Propensity Score Techniques

Now that I have established the overall estimate of attending a NSN school on students’ progress through the math pipeline from a traditional, multilevel regression approach, I will present the results of my propensity score matching (PSM) and weighting (PSW) techniques. To briefly review my methodology, I conduct three separate propensity score techniques to ensure that my estimates are robust across model specifications. First I generate propensity scores through a multilevel logistic regression that utilizes both student characteristics and the attributes of their middle schools to predict students’ propensity to attend a NSN school. I then use these scores in two different ways. First, I use them to select a matched group of control cases from the group of students who do not attend a NSN school to serve as a comparison group for the students in NSN schools\(^{13}\). I then use further regression adjustment, controlling for all pre-treatment student-level characteristics within this matched sample. I refer to this method as multilevel propensity score matching. Next, I use the propensity scores as modeling weights in a regression that includes all students in my analytic sample. Essentially, I weight the control units in such a way that they reflect the same propensity to attend a NSN school as the treatment group (see Chapter 4 for more details). I refer to this strategy as multilevel propensity score weighting. Finally, in my preferred approach, I regenerate students’ propensity scores using only student characteristics, and then force an exact match on students’ middle schools when selecting my matched sample of controls. I then perform additional regression adjustment within this matched sample. This technique essentially compares students who attended the same middle schools, but some went on to attend a NSN high school and some went on to attend an alternate high school. Because my multilevel regression results suggest differential effects of attending a NSN school by student cohort, I perform my propensity

\(^{13}\) As described in Chapter 4, I use one-to-one, nearest neighbor matching, without replacement and I restrict analyses to areas of common support. As a robustness check I also test matching with replacement and find comparable results.
score techniques separately for the 2005-06 and the 2006-07 cohorts. For all techniques, my estimates are for the treatment effect on the treated (i.e., for the students in NSN schools).

My results from the 2005-06 cohort are presented in Table 6.5 and my results from the 2006-07 cohort are presented in Table 6.6. For both cohorts, the results across my three propensity score techniques are very consistent with one another. However, estimates do differ for the two cohorts. For the students entering high school in 2005-06, I find a small positive effect of attending a NSN school. Even after restricting comparisons to students who attended the same middle school but went on to attend NSN and non-NSN schools, I find that students attending the NSN schools progress roughly one-sixth of a year farther through the math curriculum than they would have had they attended an alternate high school option. This is in direct contrast to the significant, negative effect estimated by my multilevel regression analyses. This indicates that the standard regression approach was insufficient to adjust for the bias associated with student selection into schools.

Research on small reform in the New York City context that does not go beyond standard regression risks unfairly penalizing small schools for their more disadvantaged student bodies.

**Table 6.5: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2005-06 cohort**

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multilevel PSM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Multilevel PSW&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PSM with exact match on middle school&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>N=10,434</td>
<td>N=42,172</td>
<td>N=10,100</td>
</tr>
<tr>
<td>New, Small, Nonselective (NSN) School</td>
<td>.174**</td>
<td>.178**</td>
</tr>
</tbody>
</table>

<sup>a</sup> The most advanced math course passed in four years of high school (0=No math, through 8=Advanced topics beyond Calculus)
<sup>b</sup> PSM is an abbreviation for propensity score matching, PSW is an abbreviation for propensity score weighting

For the 2006-07 cohort, however, I find no impact of attending a NSN school in any of my propensity score matching techniques. Given that the differential cohort effect of attending a NSN school remains after applying the propensity score matching techniques it is less likely that this
finding is caused by selection bias. One possible explanation for this result is that the newly created schools, that served their first cohort of ninth graders in the 06-07 school year, may have had a negative influence on student course-taking, which is cancelling out the continued positive effects of the original set of schools. To test this possibility, I re-analyze all propensity score matching strategies on a sample of students from the 06-07 cohort, which excludes students attending the newly opened schools (see Table A.11 in Appendix A). I find very similar results, indicating that the newly created schools are not the cause of the differential effects between cohorts.

Another explanation is that the initial burst of energy educators experience in starting a new school may fade over time. Moreover, since new schools in New York City were designed to phase in one grade at a time, educators’ and staff would have been able to give their full, undivided attention to the first cohort of students but would have necessarily had to spread their attention around more as each new grade was added until the schools reached full capacity. If this differential impact on course-taking between the 2005-06 and 2006-7 school years is part of larger trend of NSN schools losing their positive influence over time then it has important implications for the overall efficacy and appeal of the reform. This finding is consistent with the Hemphill & Nauer (2009) study of small high school reform in New York City, which found that the positive effects of the new, small schools on attendance and graduation rates were fading over time.

Table 6.6: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multilevel PSM(b)</td>
<td>Multilevel PSW(b)</td>
<td>PSM with exact match on middle school(b)</td>
</tr>
<tr>
<td>N=12,858</td>
<td>N=46,069</td>
<td>N=10,493</td>
</tr>
</tbody>
</table>

| New, Small, Nonselective (NSN) School | .073 | .076 | .069 |

\(^a^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)

\(^b\) PSM is an abbreviation for propensity score matching, PSW is an abbreviation for propensity score weighting
As discussed previously, extant literature indicates that the new, small high schools in New York City improve their students’ rate of graduation. My preferred analytic approach consequently retains drop-outs in my sample to avoid punishing small schools for keeping more students in school. However the result of this decision is that the NSN effect on students’ progression through the math pipeline may partially flow from the fact that they are simply preventing more students from dropping out, and not from any real impact they are having on student course-taking. To test this possibility, I re-analyze each of my propensity score matching techniques for both cohorts on a sample of students that excludes drop-outs (see Tables A.12 and A.13 in Appendix A). I find that the NSN school effects are slightly reduced across all techniques and both cohorts, but that they retain the same significance levels and substantive interpretations. For example, the NSN effect on students’ progression through the math pipeline for the 2005-06 cohort, utilizing the PSM technique with an exact match on students’ middle schools, was .175 for the sample including drop-outs and .156 for the sample excluding drop-outs. These results suggest that a small part of the NSN effect on student course-taking flows from the fact that they retain more students in school through graduation, but that this phenomena does not explain the entire effect.

**Assessing Curricular Rigor Across Schools**

In the final piece of my analyses, I assess whether the rigor of math courses varies across schools. In particular I am interested in how the standards of performance of courses in NSN schools compare to the other school options in New York City. Up until this point, all of my analyses have assumed that the most advanced math course that students successfully complete is a good indicator of their level of math content mastery. The analyses have also implicitly assumed that the math pipeline measure is an *equally good* indicator of students’ math proficiency across all schools in New York City. Moreover, recall that the pipeline measure was calculated from students’ transcript data and thus uses course titles as the primary designation for course content. However, as
discussed in more depth in Chapter 3, we know from the extant literature applying institutional theory to educational settings, that schools enrolling substantial proportions of disadvantaged and academically underprepared students may have incentives to project the image of a more advanced math curriculum than they actually have in place. In order to diagnose whether this type of “course label inflation” is occurring in New York City, and specifically in the NSN schools, I compare students’ marks in their courses to their marks on a more objective assessment—New York state Regents exams. Specifically, I subtract the average mark students receive on the courses (conditional on passing) covered by each Regents exam from their score on that Regents exam. I have chosen to call the resulting difference score the “performance differential”.

I examine the curricular rigor for course sequences leading up to both the Math A and Math B Regents exams. From an analytic standpoint, both exams have strengths and weaknesses. Many more students take the Math A exam, covering roughly Algebra I and the first semester of Geometry, because it is required for graduation. Consequently the sample size of students with Math A Regents exam scores is much more representative of the overall population of New York City students. On the other hand, because the content areas covered by the Math A exam are so basic, many of the more advanced students take the first, and often second, courses of the Math A sequence in middle school. Because it is only inappropriate to hold high schools accountable for their own courses, I only included courses taken during high school in the average course mark. This means that for high achieving students, the average course mark may only capture the last course or last two courses in their Math A sequence.

The Math B Regents exam on the other hand, which covers roughly the second semester of Geometry as well as Algebra II/Trig, is required only for students who want an Advanced Regents Diploma. As a result, many fewer students attempt the Math B Regents than the Math A Regents and those students that do attempt Math B are the more advanced students. 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 fully attributed to their high schools. Moreover, because Math B both covers more advanced content than Math A and is a somewhat lower stakes exam (students do not need to pass it to graduate), one might expect schools to be more likely to practice “course label inflation” for Math B than Math A.

Table 6.7 presents the psychometric properties of the Math A and Math B exams as well as the performance differentials for both subjects. The intraclass correlation (ICC) represents the proportion of the variability in a given outcome that is systematically between schools. For Math A, 32.6 percent of the variation in Regents scores and 26.8 percent of the variation in the performance differential occurs between schools. For Math B, 41.7 percent of the variation in Regents scores and 53.3 percent of the performance differential lies between schools. These ICCs, particularly for Math B, are relatively high for academic outcomes, suggesting the likely presence of school effects. As predicted, the between-school portion of the variance in performance differential is substantially higher for Math B than it is for Math A.

<table>
<thead>
<tr>
<th>Table 6.7: Multilevel Psychometric Properties of Math A and Math B Regents Scores and Performance Differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math A Regents Score</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>N=78,343 students in 340 schools</td>
</tr>
<tr>
<td>Pooled within-school variance ($\sigma^2$)</td>
</tr>
<tr>
<td>Between-school variance ($\tau$)</td>
</tr>
<tr>
<td>Intraclass Correlation (ICC)$^b$</td>
</tr>
</tbody>
</table>

$^a$ Performance differential refers to the gap between students’ average course marks and their Regents Exam marks for the same subject. It is calculated by subtracting the average course mark from the Regents mark.

$^b$ ICC = $\tau / (\tau + \sigma^2)$

**Performance differentials across school type.**

In Table 6.8 I present descriptive statistics on average course marks, Regents exam marks, and the performance differentials for both Math A and Math B curricular sequences differentiated by school type. I also include the percent of students that attempt each Regents exam. As discussed,
many more students attempt the Math A exam than the Math B exam. In NSN schools, 78.6 percent of students attempt Math A while only 11.3 percent even attempt Math B. This discrepancy is not all together surprising, recall from Chapter 5 that the average student in a NSN school does not make it all of the way through the Math B course sequence (equivalent to completing the second semester of Algebra II/Trigonometry). For Math A the percentage of students in NSN schools attempting the exam is on par with the district-wide average, however for Math B, students in NSN schools are among the least likely to attempt the exam (only VSN schools have a lower rate). Students in specialized high schools are well over six-times more likely to attempt the Math B Regents than students in NSN schools. Even students in the large nonselective schools are somewhat more likely to take the Math B exam (16.9 percent versus 11.3 percent).

But while important, these statistics simply reinforce the findings already presented in Chapter 5—that, unadjusted, students in NSN school progress less far through the math curriculum than students in most other school types. The purpose of these analyses however, is to assess the rigor of the courses that students do take. Also displayed in Table 6.7 are students’ average course marks for both the Math A and Math B course sequences. Interestingly, these marks do not vary widely across schools. Math A average marks range from a low of 75.7 in NSN and VSN schools to a high of 85.6 in specialized schools. Math B average marks vary even less, ranging from a low of 80.5 in midsize selective schools to a high of 86.1 in specialized schools. Looking at course grades alone, one might assume that—at least among the students that take and pass their courses—the level of content mastery does not vary greatly across school types.

The Regents exam scores paint a somewhat different picture. Average student scores on the Math A exam range from a low of 61.9 in VSN schools to a high of 90.7 in specialized high schools. The NSN schools rank near the bottom, with an average score of 63.5—high enough to graduate with a local diploma (cut-off of 55), but not a Regents diploma (cut-off of 65), and well below the
college readiness benchmark (cut-off of 80). The disparities in Math B Regents scores are even more severe, average scores range from a low of 46.8 in VSN schools to a high of 81.0 in specialized high schools. Again, NSN schools rank second from the bottom, with an average score of 50.4.

Given the relative parity of average course marks and the huge discrepancies in Regents exam marks for the same subjects, it is unsurprising that I find substantial differences in the performance differentials across schools. Recall that the performance differential is calculated by subtracting the average course mark from the relevant Regents exam mark. Thus a negative differential means that students performed worse on the Regents exam than in the course, a zero differential means they performed exactly the same on both, and a positive differential means that they performed better on the exam than in their courses. For Math A, the average performance differential across all schools was -8.86 and it ranged from -12.3 in VSN schools to +5.93 in specialized high schools. NSN schools, again second from the bottom, had an average differential of -11.4. For Math B the disparities across schools are much more extreme. Clearly a more difficult exam, the citywide average performance differential for Math B was -25.7, ranging from a low of -34.7 in VSN schools (followed closely by NSN schools with -32.0) to a high of -4.80 in specialized high schools. Across all school types students performed worse, on average, on the Math B regents exam than in their Math B course sequence. The size of the differential, however, varies greatly across schools, and at first glance appears to be associated with both school size and selectivity. Small, nonselective schools—both new and veteran— have the largest negative performance differentials in both subjects suggesting that their courses are the least rigorous (or at a minimum, do not prepare their student for the Regents exam as well as similar courses in other schools). Only 11 percent of students in NSN schools even attempted to take the Math B Regents exam and, of that already select group, the average student (who would have earned a B- average in her courses) would have failed the Math B exam.
Table 6.8: Average Course Marks, Regents Marks, and Performance Differentials of DOE High Schools Serving the 2005-2006 and 2006-07 Freshman Cohorts by School Size and Selectivity

<table>
<thead>
<tr>
<th>Course Mark</th>
<th>New, Small, Nonselective (N=116)</th>
<th>Veteran, Small, Nonselective (N=14)</th>
<th>Small, Selective (N=93)</th>
<th>Midsize, Selective (N=46)</th>
<th>Large, Nonselective (N=44)</th>
<th>Large, Selective (N=25)</th>
<th>Specialized (N=9)</th>
<th>All Schools (N=342)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Math A</td>
<td>75.7</td>
<td>75.7</td>
<td>77.4</td>
<td>77.0</td>
<td>76.2</td>
<td>77.6</td>
<td>85.8</td>
<td>76.5</td>
</tr>
<tr>
<td>Avg. Math B</td>
<td>82.6</td>
<td>82.2</td>
<td>82.7</td>
<td>80.5</td>
<td>80.8</td>
<td>81.1</td>
<td>86.1</td>
<td>81.9</td>
</tr>
<tr>
<td>Attempted Regents (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math A</td>
<td>78.6</td>
<td>62.6</td>
<td>78.4</td>
<td>82.1</td>
<td>73.4</td>
<td>82.1</td>
<td>87.5</td>
<td>78.0</td>
</tr>
<tr>
<td>Math B</td>
<td>11.3</td>
<td>8.14</td>
<td>19.3</td>
<td>26.7</td>
<td>16.9</td>
<td>27.6</td>
<td>71.1</td>
<td>17.4</td>
</tr>
<tr>
<td>Regents Mark (1st attempt)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Math A</td>
<td>63.5</td>
<td>61.9</td>
<td>69.9</td>
<td>69.6</td>
<td>67.6</td>
<td>73.3</td>
<td>90.7</td>
<td>67.3</td>
</tr>
<tr>
<td>Avg. Math B</td>
<td>50.4</td>
<td>46.8</td>
<td>57.6</td>
<td>57.2</td>
<td>61.7</td>
<td>60.9</td>
<td>81.0</td>
<td>55.9</td>
</tr>
<tr>
<td>Performance Differential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Math A</td>
<td>-11.4</td>
<td>-12.3</td>
<td>-7.71</td>
<td>-7.06</td>
<td>-8.16</td>
<td>-3.80</td>
<td>5.93</td>
<td>-8.86</td>
</tr>
<tr>
<td>Avg. Math B</td>
<td>-32.0</td>
<td>-34.7</td>
<td>-24.5</td>
<td>-23.1</td>
<td>-18.8</td>
<td>-20.0</td>
<td>-4.80</td>
<td>-25.7</td>
</tr>
</tbody>
</table>

\(^a^\) See Measures section for descriptions of school size and selectivity classifications.

\(^b^\) Column school counts do not sum to 322 because Specialized high schools are classified according to their size and selectivity as well as being separated out into a separate column.

\(^c^\) Course marks represent the average of all courses, conditional on passing, that students take during high school, before their first attempt at the Regents exam covering those courses.

\(^d^\) Performance differential refers to the gap between students' average course marks and their Regents Exam marks for the same subject. It is calculated by subtracting the average course mark from the Regents mark.
School factors associated with curricular rigor

Next, I examine which school factors are associated with students’ performance differentials, after controlling for all student-level background characteristics (see Table 6.9). Models 1 and 4 present the associations between school type and performance differentials, adjusted for all student-level covariates but not yet adjusted for student body composition or schools’ curricular structures. For Math A, I find that students in small schools have 1.3 points larger negative performance differentials, on average, than their counterparts in midsize schools, even after controlling for all student-level covariates. Students in large schools, on the other hand had 2.3 points smaller negative performance differentials, on average, than students in midsize schools. A similar pattern holds for Math B, although the coefficient for small school size is non-significant (due to a smaller sample size) it is roughly the same magnitude and direction as for Math A. Again, students in large schools have smaller negative performance differentials (4.87 points) compared to their peers in midsize schools. For both Math A and Math B, students in academically nonselective schools have larger negative performance differentials than their peers in selective schools (-1.45 and -3.21 points, respectively). Attending a new school was not related to students’ performance differentials in either subject.

Essentially, before adjusting for student body composition and curricular structures it appears that both school size and selectivity are associated with the rigor of math courses. Students’ marks in math courses in large schools and in academically selective schools are better indicators of how well they will perform on the relevant Regents exams than students’ marks in math courses in small schools and non-selective schools, even after controlling for their own entering math achievement, race/ethnicity, poverty status, and all other student-level covariates. Why might this be the case? Is it simply because smaller, nonselective schools enroll more disadvantaged students? Or does it have to do with the curricular structures these schools put in place? In the next two models I
control for both of these aspects of schooling to get a sense of what, if anything, might be mediating the relationships between school size and selectivity and curricular rigor.

In Models 2 and 5 of Table 6.9 I further adjust the school type estimates for schools’ student body composition. After controlling for the percent of black and Hispanic students as well as the average entering math achievement, the association between school selectivity and students’ performance differentials becomes non-significant for both Math A and Math B. For Math A, the school size estimates are reduced and for Math B they too become non-significant. In other words, it appears that the association between school selectivity and curricular rigor was, unsurprisingly, a function of student body composition. Meanwhile, I find that the percent of black and Hispanic students in a school is related to students’ performance differentials for both Math A and Math B. A 30 percent increase in the percent of black and Hispanic students in a school is associated with 1.14 point increase in the negative performance differential in Math A and a 4.26 point increase in Math B. In other words, schools that are more heavily black and Hispanic are also likely to have less rigorous math courses.

My results also suggest that the schools’ average entering math achievement is related to students’ performance differentials in both subjects. A one standard deviation increase in average entering math achievement is associated with a 2.44 point decrease in the negative performance differential for Math A and a 2.82 decrease in Math B. Put another way, the courses in schools where students have high average entering math achievement are likely to be more rigorous than the courses in schools with low average entering achievement. Both my findings on the percent of black and Hispanic students and on students’ average entering math ability are consistent with the literature on course label inflation, which predicts that schools that enroll more disadvantaged and academically underprepared students are more likely to water down the curricular content and lower the performance standards of courses, while retaining the same course title.
Finally, in Models 3 and 6, I control for schools’ curricular structure. I test all of my curricular structure measures but remove the predictors that are non-significant for both subjects in the interest of model parsimony. The result is that for Math A I find one significant curricular structure factor—whether schools offer AP/IB math. For Math B, both offering AP/IB math and the percent of courses offered below Algebra in difficulty are significant. After adjusting for these measures, the remaining associations between school size and curricular rigor disappear. Students attending schools that do offer AP/IB math have a smaller negative performance differential by 1.58 points on average, in Math A and a full 6.11 points in Math B, even after adjusting for all student and school characteristics. Meanwhile for Math B, a 10 percent increase in the number of courses offered below Algebra is associated with approximately a one-point decrease in the negative performance differential. Essentially, schools that offer advanced placement courses and that have fewer low-level course options have more rigorous courses overall. It is unlikely that offering an AP calculus or statistics course causes the rigor of Math A or Math B course to go up, however it is possible that these are both signs of an underlying curricular mission of offering rigorous, college preparatory math courses.
Table 6.9: Between-school, Multilevel Models Exploring the Relationships Between School Factors and Students’ Performance Differentials (2005-06 and 2006-07 cohorts)

<table>
<thead>
<tr>
<th>Math A</th>
<th>Math B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong>&lt;sup&gt;b&lt;/sup&gt; School Type</td>
<td><strong>Model 2</strong>&lt;sup&gt;b&lt;/sup&gt; + Student Body Composition</td>
</tr>
<tr>
<td>Small</td>
<td>-1.31*&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Large</td>
<td>2.30**&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Academically nonselective</td>
<td>-1.45**&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>New school</td>
<td>0.80&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

**School Composition**

- Percent Black and Hispanic: -0.038**<sup>c</sup> | -0.032**<sup>c</sup> | -0.142***<sup>c</sup> | -0.121***<sup>c</sup> |
- Avg. entering math achievement: 2.44*** | 2.18*** | 2.82** | 1.60<sup>c</sup> |

**Curricular Structure**

- School offers AP/IB math: 1.58*** | 6.11***
- Percent of courses below Algebra in difficulty: -0.002<sup>c</sup> | -0.102*<sup>c</sup> |
- Constant: -6.85*** | -6.35*** | -6.30*** | -25.1*** | -24.0*** | -24.0***

*<sup>p</sup><.05, **<sup>p</sup><.01, ***<sup>p</sup><.001
1 Performance differential refers to the gap between students’ average course marks and their Regents Exam marks for the same subject. It is calculated by subtracting the average course mark from the Regents mark.
2 Model is also adjusted for all student characteristics included in Table 6.2, Model 2. The following non-significant school-level predictors were removed in the interest of model parsimony: borough indicators, the percent of students that were free or reduced price eligible, special education, overage, ELL, and the percent of total courses classified as prep/support.
3 School size categories are compared to midsize schools.

Curricular rigor in the new, small, nonselective schools

Now that I have described which characteristics of schools are associated with the rigor of math courses, I focus in on estimating the curricular rigor for NSN schools specifically. In models 1 and 3 of Table 6.9 I establish the unadjusted relationship between attending a NSN school and students’ performance differentials in Math A and Math B. I find that, on average, students in NSN schools have roughly four-point larger negative performance differentials in Math A and nine-point larger negative performance differentials in Math B than students in alternate high school options in
New York City. Next, in Models 2 and 4, I control for all student-level covariates. Even after adjusting for student characteristics, my results suggest that students in NSN schools have a 2.56 point larger negative performance differential in Math A and a 7.27 point larger negative performance differential in Math B than their peers in other types of high schools. Put another way, the math courses in NSN schools appear to have weak curricular rigor as measured by their standards of performance. Students’ average marks in Math A and B courses in NSN schools are not as good indicators of how well they will fare on their Regents exams as they are in the alternate school options in the city. Moreover the rigor of the more advanced Math B courses appears to be weaker than that of the more basic Math A courses.

As I will discuss in greater detail in Chapter 7, these results have important implications for my assessment of the NSN school impact on progress through the math curriculum as well as for the broader assessment of small high school reform in New York City. Earlier in this chapter I presented findings indicating that the NSN schools did improve students’ progress through the math curriculum slightly for the 2005-06 cohort and had no effect for the 2006-07 cohort. These findings assumed, however, that passing a given math course was a good indicator of students’ content mastery in that subject. Moreover, they assumed that progressing far enough through the math pipeline (at least through Algebra II/Trigonometry) was a reasonable gauge of students’ college and career readiness in math. If however the rigor of courses in NSN schools is weaker than the rigor of courses covering the same content in other school types, as my results suggest, then I may have overstated the curricular benefits of attending a NSN school.
Table 6.10: Between-school, Multilevel Models Exploring the Relationship Between Attending a New, Small, Nonselective School and Students’ Performance Differentials* (2005-06 and 2006-07 cohorts)

<table>
<thead>
<tr>
<th></th>
<th>Math A</th>
<th>Math B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1 NSN school</td>
<td>Model 2 +Student Characteristics</td>
</tr>
<tr>
<td>New, small, nonselective (NSN) school</td>
<td>-3.95***</td>
<td>-2.56***</td>
</tr>
<tr>
<td><strong>Student Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-0.16</td>
<td>0.42</td>
</tr>
<tr>
<td>Black</td>
<td>-0.47**</td>
<td>-2.05***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.48**</td>
<td>-1.91***</td>
</tr>
<tr>
<td>Native/Multi</td>
<td>-1.15**</td>
<td>-0.46</td>
</tr>
<tr>
<td>Female</td>
<td>-2.29***</td>
<td>-1.91***</td>
</tr>
<tr>
<td>Special Education</td>
<td>-6.90***</td>
<td>-2.55***</td>
</tr>
<tr>
<td>English Language Learner</td>
<td>-1.74***</td>
<td>0.66*</td>
</tr>
<tr>
<td>Free or Reduced Price Lunch eligible</td>
<td>0.00</td>
<td>-0.16</td>
</tr>
<tr>
<td>Overage for 8th grade</td>
<td>-1.32***</td>
<td>0.32</td>
</tr>
<tr>
<td>Entering math achievement**</td>
<td>3.44***</td>
<td>4.13***</td>
</tr>
<tr>
<td>Entering ELA achievement**</td>
<td>1.53***</td>
<td>1.06***</td>
</tr>
<tr>
<td>Days absent in 8th grade</td>
<td>0.02***</td>
<td>-0.03*</td>
</tr>
<tr>
<td>Days tardy in 8th grade</td>
<td>0.03***</td>
<td>0.01</td>
</tr>
<tr>
<td>2006-07 Cohort</td>
<td>-0.20**</td>
<td>-0.21</td>
</tr>
<tr>
<td>Constant</td>
<td>-7.87***</td>
<td>-7.41***</td>
</tr>
</tbody>
</table>

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

* Performance differential refers to the gap between students’ average course marks and their Regents Exam marks for the same subject. It is calculated by subtracting the average course mark from the Regents mark.

** All racial/ethnic groups are compared to whites. Asian and Native American/Multi-ethnic indicator variables and interactions are included in the models to preserve white students as the comparison group but are omitted from the table because they are non-significant.

*** Measure is z-scored (M=0; SD=1)

**Summary**

The results of my analyses on NSN schools are mixed. For the first cohort of students in my analyses (2005-06), I find a small positive effect of attending a new, small school on students’ progress through the math pipeline. By the second cohort (2006-07), however, I find no effect in either direction. Moreover, my multilevel regression results suggest that NSN schools are not equally
beneficial for all types of students. Black and Hispanic students appear to do better in NSN schools than alternate high schools, while the reverse is true for whites. Along the same lines, students with initially low math achievement appear to do better in NSN schools, while students with middle-to-high levels of initial math achievement are better served in non-NSN schools. Finally, a comparison of students’ performance in math courses and their performance on the Regents exams covering the same subjects revealed that the performance standards, or “rigor” of courses in the NSN schools is lower than in alternate high school options in New York City. This finding casts what was already a small positive effect (one-sixth of a year farther through the math curriculum) for only one cohort into question. If we care about content mastery and not just getting transcript credit for specific courses then the negative result regarding rigor in NSN schools may outweigh the slight positive result regarding progress through the math pipeline. In the next and final chapter I will reflect more deeply on both the conceptual and policy implications of my results.
Chapter 7
Conclusion & Discussion

Conclusion

On the scale of possible education policy interventions, closing schools that are deemed to be failing and starting over from scratch with new ones is extreme. It involves a great deal of upheaval, a great deal of controversy, and a great deal of money. Yet between 2002 and 2010, the New York City Department of Education closed 27 large high schools, replacing them with over 200 new, small schools—and they aren’t finished. The confidence needed to carry out such an aggressive policy reform was grounded in a firm belief that the new, small high schools would provide students with both less anonymous and more rigorous educational experiences than the large schools had been capable of. Proponents of small school reform promised that these schools, which targeted underperforming students living in the most disadvantaged areas of the five boroughs, would produce not only more graduates, but also more college and career-ready graduates.

After a decade of implementation, the New York City Chancellor of schools proclaimed that both of these objectives had been met, and outlined a plan to “aggressively” continue pursuing small high school reform (Walcott, 2012).

When it comes to the city’s first objective of small high school reform—graduating more students—the Chancellor is correct. A rigorous evaluation of the new small schools confirms that they have raised graduation rates across all categories of students (Bloom et al., 2010; 2012). But when it comes to the second objective—graduating more college and career ready students—I argue that the Chancellor is on less solid ground. Through the lens of student course-taking and schools’ curricular rigor, the results from this dissertation paint a more complicated and less optimistic picture of the new, small high schools’ role in fostering college and career readiness in mathematics. As such, my findings are consistent with the extant research on small school reform in NYC, which
finds a small positive impact of the new small schools on college readiness in English, but none in math. Given both the national imperative of improving STEM education in order to build a strong workforce in these fields and the individual significance of advanced high school math courses, which make students more attractive to four year colleges, reduce the need for remediation once in college, and prepare students for higher education trajectories into sectors of the labor market characterized by growth, stability, and above-average wages—it is crucial to understand exactly how small high school reform is helping or hindering (or perhaps a bit of both) students experiences in the math curriculum. Thus, in this final chapter, I will summarize and reflect upon the results of my dissertation and their implications for small high school reform policy in New York City and nationally.

**Academic stratification**

Although my dissertation focuses most intensely on the new, small, nonselective (NSN) schools created over the past decade of high school reform in New York City, it is also the first study to descriptively look at the patterns of schools’ curriculum structures, advanced course offerings, and average student course-taking at a system-wide level. Indeed, one of the most disturbing findings of this dissertation is the intensity of academic stratification—both in terms of inputs and outcomes—across high schools in New York City. Moreover, the NSN schools, which were targeted to serve the same students as the closed, failing high schools, are situated at the bottom of this stratified system.

With regards to inputs, the new small schools enroll students who are, as a whole, more disadvantaged than average on nearly every indicator. NSN students are more heavily black and Hispanic, lower-income, lower-achieving, more likely to be overage for grade and have higher rates of eighth grade absences and tardies than the majority of school options in New York City. Only the veteran, small, nonselective schools enroll an arguably more underserved student body. Moreover, as
the new small schools age, they too are becoming even more disadvantaged, enrolling even less academically prepared students as well as more English Language Learners and special education students (as the two-year start-up exemptions from serving these categories of students expire). At the other end of the spectrum the highest achieving and most advantaged students, who are disproportionately white and Asian, sequester themselves in a handful of elite, specialized high schools, which admit students based on their scores on the supplementary Specialized High School Admissions Test or an audition.

When it comes to outcomes, the average student in New York City public schools does not complete the second semester of Algebra II/Trigonometry and the average student in any size of nonselective school does not even complete the first semester. NSN schools do outperform large nonselective schools, with regards to average student progression through the math pipeline by roughly half a semester, despite serving more disadvantaged students—yet they still fall short of completing the first semester of Algebra II/Trig. Meanwhile, students in the specialized high schools, on average, progress through the first semester of Calculus. These differences are not trivial. Extant literature suggests that progressing past Algebra II is a critical determinant of post-secondary participation and success (Adelman, 2006). Even the NYC DOE considers only math courses that are at the level of Algebra II or more advanced to be “college preparatory” when giving schools credit for promoting college and career readiness (NYC DOE, 2011a). The fact that the average student in a NSN school does not successfully complete any college preparatory math courses, even by the city’s own designations, is worrisome.

**Opportunity to learn**

Beyond average student progression through the math curriculum, it is also important to consider which courses students have access to. One of the principle criticisms of small high school reform in New York City was that the new schools would lack the capacity to offer advanced
courses and thus cut their students off from the opportunity to learn these subjects. The results from my dissertation largely vindicate these critiques. Out of all of my measures of schools’ curricular structures, the most stark differences between small and large schools were in the availability of Calculus or any Advanced Placement (AP) or International Baccalaureate (IB) math courses. Only one-fifth of NSN schools offered either of these course options, while over 90 percent of large nonselective high schools and 100 percent of specialized high schools offered both.

When weighing the pros and cons of very small high schools relative to midsize and large ones, a lack of access to advanced curricular opportunities should be at the top of the list. However it is important to note that in the current technological age, there are options beyond increasing school size to give students access to courses not offered by their school. New York City is currently piloting a program called iLearnNYC, a virtual learning program that allows students to take Advanced Placement courses and recover missed credits online (using certified NYC public school teachers), as well as experiment with “blended” learning, which combines face-to-face and online platforms. This pilot began in 2010, and was thus not available to the cohorts studied in my dissertation. Theoretically, this type of virtual learning program could provide high achieving students in the new small high schools the opportunity to take advanced math courses not offered by their schools. Whether these online courses are as effective as traditional courses and whether disadvantaged students enroll in them at similar rates as traditional courses has not yet been determined. New York City has commissioned evaluations of the iLearnNYC pilot but the findings from these studies have not yet been published. Although the results of this particular program are not yet know, I propose in my policy recommendations below that strategies such as virtual courses or the sharing of courses between small schools housed in the same buildings may provide avenues for offering advanced curricular opportunities to students attending schools that lack the capacity to offer them independently.
Equity vs. excellence

In an ideal world, schools would both be academically excellent and internally equitable with regards to the social distribution of achievement along student background characteristics such as race, class, and gender. Indeed, proponents of small high school reform believed that both equity and excellence would be easier to achieve in smaller schools. My dissertation confirms that small high schools are more internally equitable than large high schools, but I do not find them to be more academically excellent in the New York City context. In terms of equity, NSN schools have less variation in student progression through the math pipeline than the larger high school options in New York City and the within-school black-white and Hispanic-white gaps in progression through the math pipeline are eliminated or reversed in NSN schools. But from an excellence perspective, the average of the most advanced math course taken is lower in NSN schools than in all other high school options except the large, nonselective schools. My propensity score matching analyses do find a small positive impact of attending a NSN school on students’ progress through the math curriculum (one-sixth of a year) for the 2005-06 cohort, but not for the 2006-07 cohort. Yet while NSN students, who are among the most disadvantaged in the city, might be doing slightly better than they would have had they attended an alternate high school option, they are still failing to complete even one semester of Algebra II/Trigonometry, on average, which is the lowest level of math course considered to be college preparatory by the district.

Meanwhile, I find that attending a NSN school is not equally beneficial for students at all levels of entering math achievement. Students with very low math skills do better in the NSN schools than they would have in alternate high school options, confirming the results of existing studies, which find that small schools are most advantageous for disadvantaged students. However the reverse is true for middle and high achieving students, who are worse off in terms of progression through the math pipeline if they attend a NSN school.
These results reveal that the new, small high schools did indeed offer a narrower curriculum than the larger high school options in the city, leading to greater within-school equity. However, instead of that curriculum being college preparatory, as the reformers had promised, the curriculum in NSN schools appears to be constrained at a relatively low level of difficulty. This explains why only the lowest achievers appear to benefit from attending NSN schools—the curriculum may be more rigorous than these students would receive if relegated to the lowest track at a larger, more differentiated school. However, the middle- to high-achievers at NSN schools appear to be “constrained” in a negative way because these schools lack the capacity to offer more advanced math courses.

Interestingly, my results, which suggest that NSN schools are more internally equitable, but not necessarily more excellent than alternate high school options in New York City, are consistent with the findings of the Lee and Smith (1997) study, which was influential in motivating small high school reform to begin with. In their national study on school size, Lee & Smith found that very small schools with enrollments of fewer than 600 students were the most internally equitable but produced smaller learning gains than schools with enrollments between 600 and 900 students, which maximized both learning and equity. Moderately small high schools in the 600-900 range, they argued, were large enough to provide a solid curriculum, yet small enough to limit curricular differentiation. New York City’s small high school reform, however, purposefully created very small schools with student enrollments of roughly 550 students or fewer. In other words, despite the limitations in applying a study using national data from the 1980s to a current, active school reform, New York City might have benefited from following the recommendations from this early research more closely.

Finally, even though my results suggest that the new small high schools are more internally equitable than alternate school options in New York City, that does not mean that they have
substantially impacted the system-wide inequities in students’ academic experiences. As discussed earlier, the New York City public school system is highly stratified between schools. Consequently, reducing the within-school stratification among the schools serving the most disadvantaged students in the city cannot level the playing field between the haves and the have-nots.

Unequal standards

One of the weaknesses of my analyses assessing the relationship between school characteristics and students’ progress through the math curriculum is that they assume that course titles are an accurate representation of course content. However, we know from extant research applying an institutional theory framework to education that some schools may face institutional pressures to project the image of a more advanced curriculum than they actually have in place. Moreover, the literature suggests that schools serving high percentages of low-income, minority, and academically under-prepared students are more likely to practice this sort of “course label inflation” than schools serving more advantaged students. Given the level of stratification between schools in New York City, the possibility of unequal curricular rigor across school is a real concern.

My results assessing the difference between students’ performance in their math courses and their performance on the relevant New York state Regents exam indicate that this concern was warranted. I find that, even after controlling for all student background characteristics, attending a NSN school is associated with larger gaps between students’ marks in their courses and their marks on Regents exams, particularly for the more advanced, Math B courses. The average student in a NSN school does not make it through the full Math B course sequence. Indeed, only 11 percent of students in NSN schools even attempt the Math B Regents exam. And for that 11 percent, the average grade they received in their Math B courses was a B minus; the average mark they received on the Regents was an F (below the passing cut-off mark of 55). Meanwhile, over 70 percent of students in the specialized high schools attempt the Math B Regents, and for that 70 percent the
average course grade was a B and the average Regents mark was equivalent to a B minus (above the college-readiness cut-off mark of 80). Clearly, courses with the same title are not equally preparing students for their exams across schools of different types. My results also find both a negative association between the percent of black and Hispanic students enrolled in a school and the rigor of courses as well as a positive association between students’ average entering math achievement and the rigor of courses, even after adjusting for students’ own race and achievement.

**Contextualizing the constrained curriculum hypothesis**

The results of this study confirm an argument that I make in my conceptual framework, namely that the constrained curriculum hypothesis is overly simplistic. I agree with the basic premise that small schools lack the capacity to offer as differentiated a curriculum as large schools and that this has the potential to increase the rigor of students’ curricular experiences. However I argue that the effectiveness of narrowing the curriculum will depend on the level of difficulty at which it is targeted. Offering only low-level math courses will clearly not have the same effect as offering only advanced courses. Furthermore, my results suggest that the level at which schools decide to target their curricula depends on several mediating factors, including the characteristics of their student body (particularly students’ average entering level of achievement). Moreover, I argue that the institutional pressures that schools face in the current college-for-all political climate interact with these same student body characteristics to compel some schools to project the image of a more rigorous curriculum than they actually have in place. Finally, I contend that the level at which small schools target their curriculum—in reality, not just according to course titles—interacts with individual students’ entering achievement such that attending a small school with a narrow curriculum cannot be equally beneficial for all students.
Put another way, my results from evaluating small high school reform in New York City confirm that the original constrained curriculum hypothesis, formulated based on aggregate national patterns of the relationship between high school size and student learning, needs to be contextualized before it can accurately predict the outcome of an active, urban, small high school reform policy. In New York City, the new small high schools were targeted to serve the most disadvantaged and academically underserved students in the five boroughs. The fact that these schools are positioned at the bottom of a highly stratified educational system has important implications for how these schools will perform relative to the other high school options in city.

**Discussion**

**Limitations**

As is the case with most research relying on secondary data analyses, this study does have several important limitations, which I will outline briefly in this section.

**Strong assumptions needed for causal inference.** Concurrent with the introduction of small high school reform, the NYC DOE also drastically increased the level of student choice in the high school assignment process. In addition, New York City high schools already existed on a spectrum of academic selectivity, ranging from nonselective schools to the specialized high schools, which admit students based on their scores on a supplementary admissions test. This combination of students choosing schools and schools choosing students creates a system that maximizes the sorting of students among schools and consequently also maximizes the threat of selection bias. In this dissertation, I do my best to address this threat of bias by utilizing several propensity score matching techniques. However, because the only information I have on students’ background characteristics comes from the city’s administrative data, I have a limited number of student attributes to match on. This leaves open the potential for omitted variable bias. Thus I interpret the
results from my causal inference, propensity score matched models very cautiously and I do not interpret the results from any of my other analyses as causal.

Establishing a time trend. One of the interesting but largely unexplored findings from this study is that the effectiveness of the new small schools appears to differ by cohort. I find a small positive impact of NSN schools on students’ progression through the math curriculum for the 2005-06 cohort, but none for the 2006-07 cohort. Future research evaluating both earlier and later cohorts would be useful in order to establish whether these different results are part of a larger time trend of fading small school effects. This type of trend would not be inconsistent with extant research (Hemhill & Nauer, 2009) or anecdotal stories of educators in small schools investing extra energy for the inaugural class of students. Moreover, we know that the initial exemptions from serving English language learners and special education students expire after a school has been open for two years. Furthermore, as more large, “failing” high schools are closed and replaced by NSN schools, the control group of alternate high school options becomes relatively higher performing and the NSN schools enroll a greater and greater share of the city’s most disadvantaged students.

The historical counterfactual. Among the most important steps in conducting a policy evaluation is identifying the counterfactual—that is, what would have happened in the absence of the policy. However, I argue that in the case of small school reform in New York City, there are two plausible counterfactuals: the contemporary counterfactual, which assumes that if small high schools didn’t exist, then the students who currently attend them would have attended another available, mid-sized or large school, or; the historical counterfactual, which assumes that if the entire small school reform policy had never been implemented, and the large, “failing” schools had never been replaced by small schools, then students currently attending new small schools (or at least some of them) would have attended the large “failing” high schools. This study assesses the impact of NSN schools relative to the contemporary counterfactual, however understanding the impact of the NSN schools
relative to the closed high schools that they replaced is equally important to the overall impact of the
small high school reform policy in New York City. As such, the causal estimates from my analyses
are likely *conservative* relative to what they would be if I had compared NSN schools to the large,
closed high schools, which were selected for closure specifically because they had very low
graduation rates.

**Mixed methods preferred.** My dissertation employs large scale, quantitative data analyses
to describe trends and estimate the effect of attending a NSN school relative to the alternate high
school options in New York City. This is incredibly valuable in itself because it allows us to
understand what is going on in the aggregate in the largest public school system in the nation.
However, in an ideal world, a mixed-methods approach to examining small high school reform in
New York City would be preferable—both from an evaluation perspective and from a theoretical
perspective. For example, in-depth interviews with educators and administrators in the NSN schools
would allow for a much richer understanding of how curricular decisions get made at the school-
level and what role factors such as school size and student body composition play in those decisions.
Interviews with students and guidance counselors could provide insight into how students select
courses, and how they feel the constraint or differentiation of the curriculum in their school affects
their course-taking. Classroom observations and analyses of course materials and syllabi would help
explain the differences in curricular rigor across schools. These are just a few examples of how an
iterative, mixed methods study could build on and inform the quantitative results presented in my
dissertation to provide a fuller picture of small school reform in New York City. An important next
step in understanding the processes through which small high school reform affects students’
educational experiences is to explore—both quantitatively and qualitatively—potential mediators
between the reform and students’ educational outcomes. The small high schools created through
this reform were different in ways other than their size. Further exploration into these other school
factors may help explain *why* the reform has generated the results presented in this dissertation and what could be improved upon.

**Policy Implications**

The analyses that I have presented in this dissertation may lead readers to assume that small high school reform was not successful in New York City, or that the new, small schools are not an improvement over the large, closed schools, which they replaced. This would be a mistake. It is imperative that the results from my study are not interpreted alone, but rather that they are added to the existing evaluations of small high school reform. For example, we already know that the new, small schools have improved attendance and graduation rates. This in itself is a huge accomplishment. Moreover, my findings are not wholly negative—I do find a small positive impact of attending a small school on students’ progress through the math curriculum for one cohort. I also find that these schools are especially beneficial for the most low-achieving students. However my results should be interpreted in a way that complicates the unswervingly positive image of small school reform that is being projected by the New York City Department of Education. These schools are still very low performing and are not adequately preparing the majority of their students for college or careers. Thus, in this section I briefly outline five of the policy implications that I believe should be drawn from my results.

1. **Consider slightly larger schools.** The findings from this study indicate that the new, small high schools in New York City are internally equitable, but not necessarily high performing. As discussed above, this mixed result was predicted by existing research on school size, which recommends *moderately* small schools of 600-900 students as being the ideal size for fostering both equity and excellence. There are no guarantees that the
results from this study will translate directly to the challenging NYC context, yet if you are going to aggressively implement small high school reform, why not at least experiment with the size that research suggests is most effective?

2. **Adopt policies that allow students to take courses not offered by their schools.**

   Although I argue above that moderately small to mid-size schools might produce better student outcomes than the very small schools created by the most recent New York City small high school reform, I recognize that would be an expensive and logistically challenging solution. A less radical policy alternative, that would still improve disadvantaged students’ access to advanced math courses would be to adopt more policies that allow students to take courses that are not offered by their schools. The iLearnNYC pilot is a good example of this type of policy. The College Now program, which allows students in participating high schools to take college or high school credited courses through CUNY campuses is another program that allows students to take courses not offered by their schools. A third option would be to allow small schools that are housed in the same building (“educational campus” in DOE parlance) to pool their resources in order to offer at least one section of courses that have lower student demand, such as advanced math and science or additional foreign languages courses, and make them available to all students in the educational campus.

3. **Consider addressing academic stratification.** Many of the findings in my dissertation point to associations between the average entering math achievement of a school’s student body and their students’ curricular experiences and outcomes. If all schools in New York City followed the Education Option model of selecting students, which
attempts to create balanced schools, serving low, middle, and high achievers, I believe that there would be fewer inequities in average course-taking, access to advanced math courses, and curricular rigor across schools. Moreover, in a system with less between-school stratification, the equalizing potential of reducing differentiation within schools by constraining the curriculum would be much greater because it would actually ensure that students of different background were experiencing the same day-to-day course content and performance standards. Nonetheless, it is important to note that there are three types of stratification: within-schools, between-schools in the same district, and between-districts. This dissertation focuses on two types of stratification: within-schools and between-schools, but does not focus on between-district segregation. Scholars have long noted that elite schools of choice in large urban districts have served as private schools at the public expense. Supporters of such schools, however, argue that they keep socio-economically advantaged and politically adroit parents in the public school system and that the presence of these parents, with their resources and influence, will benefit the system as a whole.

4. **Continue to emphasize college preparatory courses.** I believe that the New York City Department of Education is on the right track by beginning to emphasize college preparatory courses, when assessing how well high schools are doing. But in addition to adding the college and career readiness section to high schools’ report cards, I believe it would also be beneficial to hold middle schools accountable for how many of their students take Algebra by the eighth grade. The odds of making it to college preparatory math courses are much greater if students enter high school already prepared to take Geometry (or at a minimum not needing to take Pre-Algebra).
5. **Address disparities in content standards across schools.** One of the most disturbing findings of this study is how the performance standards of courses with the same title differ across schools as a function of student body composition. The district already includes Regents exam performance on schools’ progress reports, which is good. However, I believe it would also be useful for teachers to have access to statistics such as the aggregate performance differential of their students’ course marks and Regents marks. This would allow them to better assess how well their standards align with the more objective Regents exam.

**Final thought: a common curriculum without common schools?**

In the introduction I quoted Jonathan Kozol, an author and outspoken critic of school segregation, cautioning that “Small, segregated, and unequal schools are only an incremental improvement over large, segregated and unequal schools. They don't address the basic issues” (Kozol, 2006). While the real story is more complex, Kozol makes a good point. It is almost impossible to defend the status quo in New York City before small school reform. The large, high schools selected for closure had high drop-out rates and were not preparing the majority of their students for college or careers. The results from my dissertation, taken in conjunction with the existing research on small high school reform in New York City, suggest that the new small schools are an improvement, particularly for the lowest achieving students. But just because a policy improves upon the status quo does not mean it is the best option. We’ve all heard the adage that the perfect should not be the enemy of the good, but it is also true that we should not be satisfied with a slight improvement, when very substantial problems still remain. I believe that disadvantaged students in the five boroughs have improved secondary school options because of small school reform, but I also believe that as long as we allow our school system to remain stratified by race,
class, and prior academic achievement, students’ curricular experiences will continue to vary widely across schools and these disparities will always do the greatest harm to disadvantaged students, who need quality public education the most.

Throughout our history, we have struggled over which elements of schooling should be common to all students. For the majority of the 20th century the consensus was that students from different backgrounds and with different aspirations should all attend the same, comprehensive high schools, but that the curriculum within those schools should be differentiated to meet the perceived needs of such a diverse group. In this context, reducing school size and constraining the curriculum would probably have gone a long way towards leveling the playing field of students’ curricular experiences. But as between-school and between-district stratification has increased and as the political consensus has moved towards one of “college-for-all”, we have increasingly sought policy reforms that can create a common curriculum for students of all backgrounds, without integrating them into the same schools. School choice policies, charter schools, and urban small high school reform all fall into this category.

Proponents of free-market policies such as school choice and charter schools often argue that they are improving educational equity by providing low-income, urban families with the same ability to choose their children’s school that wealthy parents have long enjoyed. But the reality of these policies is that they most often produce more between-school stratification by race and class than traditional assignment policies (Mickelson, Bottia & Southworth; 2008). As we have seen in New York City, closing large, underperforming high schools and replacing them with smaller schools targeted towards the same disadvantaged students and neighborhoods has a similar segregating effect. Small high school reform, which is often introduced in conjunction with broader choice-based assignment policies such as in New York City, is also often advocated as fostering educational equity. However, small school reform attempts to equalize student outcomes through
improving the academic experiences of the most disadvantaged students by reducing the size of their schools without decreasing the between-school or between-district stratification.

Small high school reform changes the organization of schools, charter school reform changes the management of schools, and broader school choice policies change the way students are assigned to schools, but none of these popular urban secondary school reforms directly address the rigor of the curriculum or students’ access to high-status knowledge. And because they all increase between school stratification by race and class, none of these policies results in students from different backgrounds experiencing the same day-to-day educational experiences. In other words, they do not actively foster a common curriculum. Is it even possible? Can we equalize the quality and rigor of students’ educational experiences by creating a common curriculum that works as effectively in schools of concentrated disadvantage and concentrated privilege? I will not say that it is impossible, but the historical record of public education in the United States and the results from this dissertation on one city’s secondary school reform both confirm that it has not happened yet.
References:

ACT. (2006). Ready for College and Ready for Work: Same or Different? Iowa City, IA.

ACT. (2007). Benefits of a High School Core Curriculum for students in urban high schools. Iowa City, IA.


Honawar, V. (2005, July 17). To maintain rigor, college board to audit all AP courses. Education Week.


CA. University of California Press.


Appendix A

Additional Tables and Figures

Figure A.1: Highest math course completed in four years (2005-06 & 2006-07 cohorts; includes dropouts)

Figure A.2: Highest math course completed in four years (2005-06 & 2006-07 cohorts; dropouts excluded)
Figure A.3: Propensity score balance diagnostics (2005-06 cohort; multilevel PSM & PSW)
Figure A.4: Propensity score balance diagnostics (2006-07 cohort; multilevel PSM & PSW)
Figure A.5: Propensity score balance diagnostics (2005-06 cohort; exact match on middle schools)
Figure A.6: Propensity score balance diagnostics (2006-07 cohort; exact match on middle schools)
### Table A.7: Socio-Demographic and Achievement-Based Characteristics of First-Time Ninth-Grade Students Enrolled in DOE High Schools by Size and Selectivity, 2006-2007 School Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49.7</td>
<td>48.5</td>
<td>43.5</td>
<td>52.0</td>
<td>50.9</td>
<td>49.4</td>
<td>51.2</td>
<td>49.7</td>
</tr>
<tr>
<td>Female</td>
<td>51.3</td>
<td>51.5</td>
<td>56.5</td>
<td>48.0</td>
<td>49.1</td>
<td>50.6</td>
<td>48.8</td>
<td>50.3</td>
</tr>
<tr>
<td>Race/Ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>3.31</td>
<td>11.3</td>
<td>9.10</td>
<td>9.06</td>
<td>18.6</td>
<td>24.9</td>
<td>50.6</td>
<td>15.1</td>
</tr>
<tr>
<td>Black</td>
<td>41.2</td>
<td>25.3</td>
<td>39.9</td>
<td>33.9</td>
<td>27.2</td>
<td>25.7</td>
<td>10.3</td>
<td>31.3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>51.0</td>
<td>57.2</td>
<td>39.8</td>
<td>43.1</td>
<td>34.4</td>
<td>36.9</td>
<td>9.84</td>
<td>39.6</td>
</tr>
<tr>
<td>Native/Multi</td>
<td>0.73</td>
<td>1.10</td>
<td>0.89</td>
<td>0.94</td>
<td>0.87</td>
<td>0.69</td>
<td>0.76</td>
<td>0.82</td>
</tr>
<tr>
<td>White</td>
<td>3.68</td>
<td>5.05</td>
<td>10.2</td>
<td>12.9</td>
<td>18.8</td>
<td>11.8</td>
<td>28.6</td>
<td>13.0</td>
</tr>
<tr>
<td>Eligible for free or reduced-price lunch (%)</td>
<td>86.5</td>
<td>88.5</td>
<td>77.7</td>
<td>78.6</td>
<td>72.8</td>
<td>76.8</td>
<td>54.6</td>
<td>77.2</td>
</tr>
<tr>
<td>Overage for 8th grade (%)</td>
<td>21.9</td>
<td>30.5</td>
<td>14.6</td>
<td>14.9</td>
<td>19.8</td>
<td>13.2</td>
<td>2.28</td>
<td>17.9</td>
</tr>
<tr>
<td>Special education (%)</td>
<td>13.3</td>
<td>10.6</td>
<td>11.7</td>
<td>11.4</td>
<td>10.8</td>
<td>8.18</td>
<td>0.64</td>
<td>10.8</td>
</tr>
<tr>
<td>English Language Learner (%)</td>
<td>15.2</td>
<td>37.2</td>
<td>8.43</td>
<td>8.74</td>
<td>18.1</td>
<td>11.2</td>
<td>2.37</td>
<td>14.6</td>
</tr>
<tr>
<td>Days absent in 8th grade (Mean)</td>
<td>18.3</td>
<td>17.0</td>
<td>14.8</td>
<td>15.3</td>
<td>17.6</td>
<td>13.0</td>
<td>7.22</td>
<td>16.1</td>
</tr>
<tr>
<td>Days tardy in 8th grade (Mean)</td>
<td>22.1</td>
<td>20.9</td>
<td>17.7</td>
<td>16.9</td>
<td>17.6</td>
<td>13.9</td>
<td>5.86</td>
<td>17.5</td>
</tr>
<tr>
<td>Entering Math achievementh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.384</td>
<td>-0.373</td>
<td>0.030</td>
<td>0.066</td>
<td>-0.100</td>
<td>0.412</td>
<td>1.60</td>
<td>0.00</td>
</tr>
<tr>
<td>SD</td>
<td>0.876</td>
<td>0.937</td>
<td>0.917</td>
<td>0.978</td>
<td>0.934</td>
<td>1.10</td>
<td>0.722</td>
<td>1.00</td>
</tr>
<tr>
<td>Entering ELA achievementh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.295</td>
<td>-0.389</td>
<td>0.078</td>
<td>0.070</td>
<td>-0.125</td>
<td>0.336</td>
<td>1.48</td>
<td>0.00</td>
</tr>
<tr>
<td>SD</td>
<td>0.843</td>
<td>0.861</td>
<td>0.963</td>
<td>1.04</td>
<td>0.928</td>
<td>1.11</td>
<td>0.860</td>
<td>1.00</td>
</tr>
</tbody>
</table>

---

*a See Measures section for descriptions of school size and selectivity classifications.

*b Column school counts do not sum to 322 because Specialized high schools are classified according to their size and selectivity as well as being separated out into a separate column.

*h Measure is z-scored (M=0; SD=1)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Entering math achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
<td>-0.402</td>
<td>-0.507</td>
<td>0.010</td>
<td>-0.007</td>
<td>-0.222</td>
<td>0.224</td>
<td>1.47</td>
<td>-0.170</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
<td>0.832</td>
<td>0.802</td>
<td>0.728</td>
<td>0.775</td>
<td>0.860</td>
<td>0.779</td>
<td>0.632</td>
<td>0.796</td>
</tr>
<tr>
<td>Entering reading achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
<td>-0.347</td>
<td>-0.486</td>
<td>0.057</td>
<td>-0.013</td>
<td>-0.241</td>
<td>0.155</td>
<td>1.39</td>
<td>-0.145</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
<td>0.798</td>
<td>0.890</td>
<td>0.784</td>
<td>0.827</td>
<td>0.856</td>
<td>0.823</td>
<td>0.817</td>
<td>0.810</td>
</tr>
<tr>
<td>Distance through the Math Curricular Pipeline (4 yr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
<td>3.23</td>
<td>3.40</td>
<td>3.78</td>
<td>3.75</td>
<td>3.14</td>
<td>3.83</td>
<td>5.69</td>
<td>3.49</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
<td>1.16</td>
<td>1.20</td>
<td>1.18</td>
<td>1.36</td>
<td>1.48</td>
<td>1.40</td>
<td>1.22</td>
<td>1.26</td>
</tr>
<tr>
<td>Distance through the Math Curricular Pipeline (1st course)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. w/in school mean</td>
<td>1.37</td>
<td>1.47</td>
<td>1.52</td>
<td>1.45</td>
<td>1.49</td>
<td>1.60</td>
<td>2.01</td>
<td>1.46</td>
</tr>
<tr>
<td>Avg. w/in school sd</td>
<td>0.275</td>
<td>0.394</td>
<td>0.378</td>
<td>0.417</td>
<td>0.468</td>
<td>0.421</td>
<td>0.532</td>
<td>0.363</td>
</tr>
<tr>
<td>School offers Calculus (%)</td>
<td>19.8</td>
<td>21.4</td>
<td>47.8</td>
<td>69.6</td>
<td>88.4</td>
<td>84.0</td>
<td>100.0</td>
<td>48.2</td>
</tr>
<tr>
<td>School offers AP/IB math (%)</td>
<td>8.6</td>
<td>21.4</td>
<td>29.3</td>
<td>58.7</td>
<td>79.1</td>
<td>68.0</td>
<td>100.0</td>
<td>35.6</td>
</tr>
<tr>
<td>Percent of credited courses below Algebra in difficulty</td>
<td>9.43</td>
<td>7.17</td>
<td>7.20</td>
<td>7.95</td>
<td>9.21</td>
<td>6.05</td>
<td>4.38</td>
<td>8.26</td>
</tr>
<tr>
<td>Percent of courses in “Prep/Support” category</td>
<td>7.13</td>
<td>5.21</td>
<td>5.57</td>
<td>5.62</td>
<td>5.68</td>
<td>3.60</td>
<td>0.27</td>
<td>5.95</td>
</tr>
</tbody>
</table>

*a* Curriculum structure group means are calculated on a sample that does not include students who transfer high schools so as not to attribute course-taking to the wrong school. Entering achievement group means are calculated on the full entering 9th grade class.

*b* See Measures section for descriptions of school size and selectivity classifications.

*c* Column school counts do not sum to 322 because Specialized high schools are classified according to their size and selectivity as well as being separated out into a separate column.

*d* Within school means and standard deviations of Eighth-grade test scores (M=0; SD=1), averaged over school size categories.

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

*f* Level of first math course taken in high school (0=No math, through 8=Advanced topics beyond Calculus)
Table A.9: Group Means of Student Characteristics, Differentiated by the Highest Math Course Completed
(2006-07 cohort; N=52,727 students in 340 schools)

<table>
<thead>
<tr>
<th></th>
<th>Below algebra II (N=19,655) Mean</th>
<th>Algebra II/Trig but below pre-calculus (N=21,294) Mean</th>
<th>Pre-calculus or above (N=11,778) Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Female</td>
<td>45.2***</td>
<td>53.1</td>
<td>55.1***</td>
</tr>
<tr>
<td>Percent Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>7.77***</td>
<td>13.0</td>
<td>35.6***</td>
</tr>
<tr>
<td>Black</td>
<td>34.0*</td>
<td>32.7</td>
<td>19.3***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>47.7***</td>
<td>39.5</td>
<td>21.9***</td>
</tr>
<tr>
<td>Native/Multi</td>
<td>0.90</td>
<td>0.83</td>
<td>0.92</td>
</tr>
<tr>
<td>White</td>
<td>9.56***</td>
<td>14.0</td>
<td>22.3***</td>
</tr>
<tr>
<td>Percent eligible for free and reduced-price lunch</td>
<td>80.1***</td>
<td>77.9</td>
<td>66.1***</td>
</tr>
<tr>
<td>Percent overage for eighth grade</td>
<td>28.7***</td>
<td>12.9</td>
<td>6.95***</td>
</tr>
<tr>
<td>Percent Special education</td>
<td>19.9***</td>
<td>7.48</td>
<td>1.41***</td>
</tr>
<tr>
<td>Percent English Language Learner</td>
<td>17.4***</td>
<td>14.3</td>
<td>9.08***</td>
</tr>
<tr>
<td>Days absent in eighth grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.6***</td>
<td>13.0</td>
<td>8.30***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(18.0)</td>
<td>(10.1)</td>
<td>(7.40)</td>
</tr>
<tr>
<td>Days tardy in eighth grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>23.5***</td>
<td>14.3</td>
<td>7.16***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(25.0)</td>
<td>(19.2)</td>
<td>(12.9)</td>
</tr>
<tr>
<td>Entering math achievement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-.617***</td>
<td>-.010</td>
<td>.968***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(.865)</td>
<td>(.744)</td>
<td>(.805)</td>
</tr>
<tr>
<td>Entering reading achievement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-.527***</td>
<td>-.012</td>
<td>.806***</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>(.860)</td>
<td>(.835)</td>
<td>(.917)</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01, ***p<.001; For significance testing, Below Algebra II (column 1) and Pre-Calculus and above (column 3) are compared to Algebra II/Trig but below Pre-Calculus (column 2).

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

bIn order to correctly attribute curricular structure measures to the appropriate schools, these measures are calculated from a student sample which includes only students who did not transfer schools part way through their high school career. Students who attend only one high school but drop out are maintained in the sample.

cMeasure is z-scored (M=0; SD=1)
Table A.10: Structural and Compositional Characteristics of High Schools, Differentiated by the School-average of Students’ Highest Math Course Completed* (2006-07 cohort; N=340)

<table>
<thead>
<tr>
<th></th>
<th>Low average course-taking(^b) (N=39)</th>
<th>Middle average course-taking(^b) (N=257)</th>
<th>High average course-taking(^b) (N=44)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School Structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Size(%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small(^c)</td>
<td>56.4</td>
<td>67.7</td>
<td>54.6</td>
</tr>
<tr>
<td>Mid-Size</td>
<td>10.2</td>
<td>15.5</td>
<td>22.2</td>
</tr>
<tr>
<td>Large</td>
<td>12.8</td>
<td>14.4</td>
<td>25.0</td>
</tr>
<tr>
<td>Specialized</td>
<td>0.0</td>
<td>0.0</td>
<td>20.5(^***)</td>
</tr>
<tr>
<td>Academically non-selective (%)</td>
<td>66.7</td>
<td>57.2</td>
<td>6.82(^***)</td>
</tr>
<tr>
<td>New School (%)</td>
<td>38.5</td>
<td>52.5</td>
<td>29.6(^*)</td>
</tr>
<tr>
<td>New<em>Small</em>Non-selective (%)</td>
<td>33.3</td>
<td>40.1</td>
<td>0.0(^***)</td>
</tr>
<tr>
<td><strong>School Composition</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Black and Hispanic</td>
<td>93.2</td>
<td>86.5</td>
<td>51.8(^***)</td>
</tr>
<tr>
<td>Percent Free or Reduced Price Lunch Eligible</td>
<td>72.4</td>
<td>70.8</td>
<td>46.5(^***)</td>
</tr>
<tr>
<td>Percent average for 8th grade</td>
<td>26.8(^***)</td>
<td>20.2</td>
<td>7.87(^***)</td>
</tr>
<tr>
<td>Percent English Language Learner</td>
<td>13.7</td>
<td>15.1</td>
<td>7.25(^*)</td>
</tr>
<tr>
<td>Percent Special Education</td>
<td>16.8(^***)</td>
<td>12.5</td>
<td>5.62(^***)</td>
</tr>
<tr>
<td>Entering math achievement(^f)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average w/in school mean</td>
<td>-.534(^***)</td>
<td>-.275</td>
<td>.747(^***)</td>
</tr>
<tr>
<td>Average w/in school std. dev.</td>
<td>.835</td>
<td>.806</td>
<td>.705(^***)</td>
</tr>
<tr>
<td>Entering reading achievement(^f)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average w/in school mean</td>
<td>-.495(^***)</td>
<td>-.238</td>
<td>.701(^***)</td>
</tr>
<tr>
<td>Average w/in school std. dev.</td>
<td>.832</td>
<td>.804</td>
<td>.826</td>
</tr>
<tr>
<td><strong>Curriculum Structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance through the Math Curricular Pipeline (4 yr)(^a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average w/in school mean</td>
<td>2.46(^***)</td>
<td>3.37</td>
<td>5.09(^***)</td>
</tr>
<tr>
<td>Average w/in school std. dev.</td>
<td>1.26</td>
<td>1.23</td>
<td>1.39(^**)</td>
</tr>
<tr>
<td>Distance through the Math Curricular Pipeline (1st course)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average w/in school mean</td>
<td>1.29</td>
<td>1.41</td>
<td>1.91(^***)</td>
</tr>
<tr>
<td>Average w/in school std. dev.</td>
<td>.344</td>
<td>.339</td>
<td>.524(^***)</td>
</tr>
<tr>
<td>School offers Calculus(^a) (%)(^***)</td>
<td>15.4(^**)</td>
<td>44.4</td>
<td>100.0(^***)</td>
</tr>
<tr>
<td>School offers AP /IB math(^a) (%)(^***)</td>
<td>12.8</td>
<td>30.4</td>
<td>86.4(^***)</td>
</tr>
<tr>
<td>Percent of credited courses below Algebra in difficulty</td>
<td>11.9(^**)</td>
<td>8.27</td>
<td>5.01(^*)</td>
</tr>
<tr>
<td>Percent of courses in “Prep/Support” category</td>
<td>8.42</td>
<td>6.15</td>
<td>2.55(^**)</td>
</tr>
</tbody>
</table>

*\(p<.05, **p<.01, ***p<.001;\) For significance testing, low average course-taking and high average course-taking groups are compared to middle average course-taking.
\(^a\) The most advanced math course passed (0=No math, through 8=Advanced topics beyond Calculus)
\(^b\) In schools with low average course-taking, the most advanced math course completed by the average student was below Algebra II in difficulty.
\(^c\) In schools with medium average course-taking, the most advanced math course completed by the average student was Algebra II (first or second term).
\(^d\) In schools with high average course-taking, the most advanced math course completed by the average student was beyond Algebra II in difficulty.
\(^e\) See Measures section for descriptions of school size and selectivity classifications.
\(^f\) Within school means and standard deviations of Eighth-grade test scores (\(M=0; \ SD=1\), averaged over school size categories.
**Table A.11: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort; Newly opened schools excluded**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multilevel PSM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Multilevel PSW&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PSM with exact match on middle school&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>New, Small, Nonselective (NSN) School</td>
<td>N=11,792</td>
<td>N=45,070</td>
<td>N=9,665</td>
</tr>
<tr>
<td></td>
<td>.076</td>
<td>.083</td>
<td>.074</td>
</tr>
</tbody>
</table>

<sup>*p<.05, **p<.01, ***p<.001</sup>

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

<sup>b</sup> For these analyses, I exclude schools which served their first cohort of 9<sup>th</sup> graders during the 2006-07 school year.

<sup>c</sup> PSM is an abbreviation for propensity score matching, PSW is an abbreviation for propensity score weighting

---

**Table A.12: Propensity Score Analyses Investigating the Effect of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2005-06 cohort; Drop-outs excluded**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multilevel PSM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Multilevel PSW&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PSM with exact match on middle school&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>New, Small, Nonselective (NSN) School</td>
<td>N=9,904</td>
<td>N=39,806</td>
<td>N=9,576</td>
</tr>
<tr>
<td></td>
<td>.141&lt;sup&gt;*&lt;/sup&gt;</td>
<td>.153&lt;sup&gt;*&lt;/sup&gt;</td>
<td>.156&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>*p<.05, **p<.01, ***p<.001</sup>

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

<sup>b</sup> PSM is an abbreviation for propensity score matching, PSW is an abbreviation for propensity score weighting

---

**Table A.13: Propensity Score Analyses Investigating the Effects of the New, Small, Nonselective Schools on Students’ Progression Through the Mathematics Pipeline, 2006-07 cohort; Drop-outs excluded**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multilevel PSM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Multilevel PSW&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PSM with exact match on middle school&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>New, Small, Nonselective (NSN) School</td>
<td>N=12,506</td>
<td>N=44,985</td>
<td>N=10,196</td>
</tr>
<tr>
<td></td>
<td>.063</td>
<td>.068</td>
<td>.060</td>
</tr>
</tbody>
</table>

<sup>*p<.05, **p<.01, ***p<.001</sup>

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

<sup>b</sup> PSM is an abbreviation for propensity score matching, PSW is an abbreviation for propensity score weighting